

41212 SEARCH REQUEST FORM

Access DB# _____

Scientific and Technical Information Center

Requester's Full Name: Natalie Davis Examiner #: 78462 Date: 7-23-01
 Art Unit: 1042 Phone Number 308-6410 Serial Number: 09/509775 CRFE
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL
8E12 CMI/9809

If more than one search is submitted, please prioritize searches in order of need.

 Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Please search claims 1-5 + 16-17

Please search SEQ ID NO: 1+2
 and

amino acids at position 14 to 226 of SEQ ID NO: 2
 amino acids at position 1 to 226 of SEQ ID NO: 2

Sept / Structure

RECEIVED
 JUL 23 2
 (STC)

STAFF USE ONLY

Searcher:	Type of Search	Vendors and cost where applicable
<u>D. Schreiber</u>	NA Sequence (#) <u>1</u>	STN _____
Searcher Phone #: <u>308-4292</u>	AA Sequence (#) <u>3</u>	Dialog _____
Searcher Location: <u>CMI 12E18</u>	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: <u>8/13</u>	Bibliographic _____	Dr. Link _____
Date Completed: <u>8/20</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: <u>8</u>	Fulltext _____	Sequence Systems <u>CompuGen</u>
Clerical Prep Time: _____	Patent Family _____	WWW/Internet _____
Online Time: <u>8</u>	Other _____	Other (specify) _____

PTO-1590 (1-2000)

BEST AVAILABLE COPY

THIS PAGE BLANK (USP)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2001, 07:46:43 ; Search time 127.64 Seconds
(without alignments) 3837.067 Million cell updates

Title: US-09-509-775-1
Perfect score: 780
Sequence: 1 tggtaagagctctaacggctg.....gaatggctgaagggttaaca 780

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database : N_Geneseq_0601.*
1: /SIDS1/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn/NA2001.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query %		Length	DB	ID	Description
		Match	Mismatch				
1	780	100.0		780	20	AA355852	DNA encoding human
2	706	90.5		1468	18	AAT66425	Human P28 coding s
3	678	86.9		678	18	AAT66424	Human P28 coding s
4	593	76.0		696	20	AA355854	DNA encoding rat g
5	584	74.9		696	20	AA355853	DNA encoding mouse
6	93.4	12.0		1056	21	AA558272	D. immitis ankyrin
7	93.4	12.0		1056	21	AA558273	D. immitis ankyrin
8	93.4	12.0		5235	19	AAV63022	D. immitis ankyrin
9	93.4	12.0		5235	19	AAV63023	D. immitis ankyrin
10	93.4	12.0		5235	21	AA558195	D. immitis ankyrin
11	93.4	12.0		5235	21	AA558196	D. immitis ankyrin

12	93.4	12.0	5503	19	AAV63314	Full length nucle
13	93.4	12.0	5503	19	AAV63020	D. immitis ankyrin
C 14	93.4	12.0	5503	19	AAV63021	D. immitis ankyrin
15	93.4	12.0	5503	21	AAAS8193	D. immitis ankyrin
C 16	93.4	12.0	5503	21	AAAS8194	D. immitis ankyrin
17	86.6	11.1	911	19	AAV63312	Nucleotide nDIank9
18	86.6	11.1	911	19	AAV63012	D. immitis ankyrin
C 19	86.6	11.1	911	19	AAV63013	D. immitis ankyrin
20	86.6	11.1	911	21	AAAS8195	D. immitis ankyrin
C 21	86.6	11.1	911	21	AAAS8196	D. immitis ankyrin
22	85.6	11.0	909	19	AAV63014	D. immitis ankyrin
C 23	85.6	11.0	909	19	AAV63015	D. immitis ankyrin
24	85.6	11.0	909	21	AAAS8197	D. immitis ankyrin
C 25	85.6	11.0	909	21	AAAS8198	D. immitis ankyrin
26	78.8	10.1	3984	22	AAFS63838	Human tankyrasel c
27	78.8	10.1	4134	21	AAAC66826	Human tankyrasel I
28	78.8	10.1	4134	21	AAAC29627	Human tankyrasel CD
29	78.8	10.1	4491	21	AAAC29632	Human truncated ta
30	78.8	10.1	4637	21	AAAC29633	Human truncated ta
31	78.6	10.1	908	19	AAV63315	Nucleotide nbmAnk9
32	78.6	10.1	908	19	AAV63024	B. malayi ankyrin
C 33	78.6	10.1	908	19	AAV63025	B. malayi ankyrin
C 34	78.6	10.1	908	21	AAAS8197	B. malaya ankyrin
C 35	78.6	10.1	908	21	AAAS8198	B. malaya ankyrin
36	77.6	9.9	906	19	AAV63026	B. malayi ankyrin
C 37	77.6	9.9	906	19	AAV63027	B. malayi ankyrin
38	77.6	9.9	906	21	AAAS8199	B. malaya ankyrin
C 39	77.6	9.9	906	21	AAAS8200	B. malaya ankyrin
C 40	74.4	9.5	573	19	AAV63311	Nucleotide nDIank5
41	74.4	9.5	573	19	AAV63010	D. immitis ankyrin
C 42	74.4	9.5	573	19	AAV63011	D. immitis ankyrin
43	74.4	9.5	573	21	AAAS8193	D. immitis ankyrin
C 44	74.4	9.5	573	21	AAAS8194	D. immitis ankyrin
45	71.6	9.2	3254	22	AAAC85295	Mouse SPANK cDNA.

ALIGNMENTS

RESULT	1	
AAX35852		
ID	AAX35852	standard; DNA; 780 BP.
XX		
AC	AAX35852;	
XX		
DT	14-JUL-1999	(first entry)
XX		
DE	DATA	encoding human gankyrin protein.
XX		
KW	Gankyrin; apoptosis induction; diagnosis; treatment; cancer;	
KW	hepatocellular carcinoma; oncogenesis mechanism; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	WO9918201-A1.	
XX		
PD	15-APR-1999.	
XX		
PF	02-OCT-1998;	98WO-JP04467.
XX		
PR	03-OCT-1997;	97JP-0286214.
XX		
PA	(FUJI/) FUJITA.	
XX		
PI	Fujita J;	
XX		
DR	WPI; 1999-277266/23.	
DR	P-PSDB; AAY02430.	
XX		
PT	Gankyrin polypeptides, useful for treatment and diagnosis of	
PT	cancers, e.g. hepatocellular carcinoma, and study of oncogen	
PT	mechanism	
XX		

```
PS Claim 5; Page 68-69; 11lpp; Japanese.
XX The specification describes human, murine and rat gankyrin DNA and
CC . polypeptide sequences. Gankyrin polypeptides inhibit tumorigenic
CC ability and apoptosis induction. The polypeptides and their antibodies
CC can be used in the diagnosis and treatment of cancers,
CC e.g. hepatocellular carcinoma, and study of oncogenesis mechanism.
CC The present sequence encodes human gankyrin.
XX
SQ Sequence 780 BP; 222 A; 161 C; 221 G; 176 T; 0 other;

Query Match 100.0%; Score 780; DB 20; Length 780;
Best Local Similarity 100.0%; Pred. No. 8.7e-243;
Matches 780; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tggtagagcttaacgctgttttgactggctagccggagccggcgacgtgagcgggc 60
DB 1 tggtagagcttaacgctgttttgactggctagccggagccggcgacgtgagcgggc 60
QY 61 gttgctgcgcgacaagtagttgctggagacagcgaatggagggtgtgtctaaacta 120
DB 61 gttgctgcgcgacaagtagttgctggagacagcgaatggagggtgtgtctaaacta 120
QY 121 atggtctgcaacctggcctacagcgggaagctggaagagttgaaggagagtattctggcc 180
DB 121 atggtctgcaacctggcctacagcgggaagctggaagagttgaaggagagtattctggcc 180
QY 181 gataaatccctggctactagaaactgacacgagcagcgaactgcaactggcgcatgc 240
DB 181 gataaatccctggctactagaaactgacacgagcagcgaactgcaactggcgcatgc 240
QY 241 tcagctgacatacagaaattgttgaattttgttgaacttgaacttggcgcagtgaaatgat 300
DB 241 tcagctgacatacagaaattgttgaattttgttgaacttgaacttggcgcagtgaaatgat 300
QY 301 aaagacgatgcaggttggtctcctctcctcaattgctgcttctgctgcccgggatgatt 360
DB 301 aaagacgatgcaggttggtctcctctcctcaattgctgcttctgctgcccgggatgatt 360
QY 361 gtaaaagccctctgggaaaaaggtgctcaagtgaatgctgtcactcaataaaatggctgact 420
DB 361 gtaaaagccctctgggaaaaaggtgctcaagtgaatgctgtcactcaataaaatggctgact 420
QY 421 cccctacattatgcagcttcgaaacagcagcagcagcagcagcagcagcagcagcagc 480
DB 421 cccctacattatgcagcttcgaaacagcagcagcagcagcagcagcagcagcagcagc 480
QY 481 ggggctaatccagatgctaaagaccattatgaggtacagcaatgcacccggcgagcagcc 540
DB 481 ggggctaatccagatgctaaagaccattatgaggtacagcaatgcacccggcgagcagcc 540
QY 541 aagggttaactgaagatgattcattacattccttctgactacaagaagcagcagcagcagc 600
DB 541 aagggttaactgaagatgattcattacattccttctgactacaagaagcagcagcagcagc 600
QY 601 gacactgaggggtaaacactcctctacacttagcctgtgtaggagagagtggaagaagca 660
DB 601 gacactgagggtaaacactcctctacacttagcctgtgtaggagagagtggaagaagca 660
QY 661 aaactgctggttcccaaggagcaagtagtttacattgagaataaagaagaagacaccc 720
DB 661 aaactgctggttcccaaggagcaagtagtttacattgagaataaagaagaagacaccc 720
QY 721 ctgcagagtgccaaaggtggcctgggttttaataactcaagagaatggtggaaggttaaca 780
DB 721 ctgcagagtgccaaaggtggcctgggttttaataactcaagagaatggtggaaggttaaca 780

RESULT 2
ID AAT66425
XX AAT66425 standard; cDNA to mRNA; 1468 BP.
```

```
AC AAT66425;
XX 17-JUN-1997 (first entry)
XX Human P28 coding sequence.
XX Human; proteasome; P28; diagnosis; malignant tumour; ds.
XX Homo sapiens.
XX Key Location/Qualifiers
FH CDS 23..703
FT /*tag= a
FT /product= p28
XX JP09075085-A.
XX 25-MAR-1997.
XX 13-SEP-1995; 95JP-0235052.
XX 13-SEP-1995; 95JP-0235052.
XX (SAGA ) SAGAMI CHEM RES CENTRE.
XX WPI; 1997-239267/22.
XX P-PSDB; AAW15483.
XX Human 26S proteasome constituting component protein - useful in the
XX diagnosis of e.g. malignant tumour
XX Claim 4; Page 7-8; 9pp; Japanese.
XX This sequence encodes the human proteasome component protein P28.
XX The protein, P28, is useful for the diagnosis and treatment of
XX various diseases caused by proteasomes such as malignant tumour.
XX
SQ Sequence 1468 BP; 424 A; 268 C; 324 G; 452 T; 0 other;

Query Match 90.5%; Score 706; DB 18; Length 1468;
Best Local Similarity 100.0%; Pred. No. 1.3e-218;
Matches 706; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 aagtagttgctggagcagcgaatggagggtgtgtgtgtctaaactaatggtctgcaacct 134
DB 1 aagtagttgctggagcagcgaatggagggtgtgtgtgtctaaactaatggtctgcaacct 60
QY 135 ggcctacagcgggaagctggaagagttgaaaggagagtagttctgcccgaataatccctggc 194
DB 61 ggcctacagcgggaagctggaagagttgaaaggagagtagttctgcccgaataatccctggc 120
QY 195 tactagaactgaccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 254
DB 121 tactagaactgaccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 180
QY 255 agaaattgttgaatttttggtaacttggagtgccagtgccagtgaaatgataaagacagcagc 314
DB 181 agaaattgttgaatttttggtaacttggagtgccagtgccagtgaaatgataaagacagcagc 240
QY 315 ttggtctcctctcaattcctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 374
DB 241 ttggtctcctctcaattcctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 300
QY 375 ggggaaaagggtgctcaagtgaatgctgcaatcaaatggctgactcccttacctatgc 434
DB 301 ggggaaaagggtgctcaagtgaatgctgcaatcaaatggctgactcccttacctatgc 360
QY 435 agcttcgaaaaaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 494
DB 361 agcttcgaaaaaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 420
QY 495 tgctaaaggaccattatgaggctacagcaatgcacccggcgagcagcagcagcagcagcagc 554
```


Db 421 tgctaaggaccattatgagctacagcaatgcacccggcagcagcaaggtaacttgaa 480
QY 555 gatgattcatacctctgtactacaagcatcccaacaacatcccaagacactgagggtaa 614
Db 481 gatgattcatacctctgtactacaagcatcccaacaacatcccaagacactgagggtaa 540
QY 615 cactcctcacacttagcctgtgatgagagagagtgaagaagcaaaactgctggtgc 674
Db 541 cactcctcacacttagcctgtgatgagagagagtgaagaagcaaaactgctggtgc 600
QY 675 ccaaggagcaagtattacattgagataaagaagaagaagacacccctgcaagtggccaa 734
Db 601 ccaaggagcaagtattacattgagataaagaagaagaagacacccctgcaagtggccaa 660
QY 735 aggtggcctgggttaatactacaagagaatggtggaaggttaaaaaa 780
Db 661 aggtggcctgggttaatactacaagagaatggtggaaggttaaaaaa 706

RESULT 3

AAAT66424 standard; cDNA to mRNA; 678 BP.

AAAT66424;

17-JUN-1997 (first entry)

Human P28 coding sequence.

Human; proteasome; P28; diagnosis; malignant tumour; ds.

Homo sapiens.

JP09075085-A.

25-MAR-1997.

13-SEP-1995; 95JP-0235052.

13-SEP-1995; 95JP-0235052.

(SAGA) SAGAMI CHEM RES CENTRE.

WPI; 1997-239267/22.

P-PSDB; AAW15483.

Human 26S proteasome constituting component protein - useful in the diagnosis of e.g. malignant tumour

Claim 1; Page 6-7; 9pp; Japanese.

This sequence encodes the human proteasome component protein P28. The protein, P28, is useful for the diagnosis and treatment of various diseases caused by proteasomes such as malignant tumour.

Sequence 678 BP; 201 A; 138 C; 183 G; 156 T; 0 other;

Query Match 86.9%; Score 678; DB 18; Length 678;
Best Local Similarity 100.0%; Pred. No. 9.7e-210;
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 atggaggggtgtgtctaaacctaatgtctgcaacctggcctacagcgggaagctgaa 156
Db 1 atggaggggtgtgtctaaacctaatgtctgcaacctggcctacagcgggaagctgaa 60
QY 157 gagttgaaggagagatctctggccgataaaatccctggctactagaaactgaccaggacagc 216
Db 61 gagttgaaggagagatctctggccgataaaatccctggctactagaaactgaccaggacagc 120
QY 217 aqaactgcattgcactggcctgctcagctgacacatacagaaattgtgaattttgtg 276

Db 121 aqaactgcattgcactggcctgctcagctgacacatacagaaattgtgaattttgtg 180
QY 277 caacttgagtgcccaagtgaatgataaagaacgatgaggttggtctcctcttcattatgcg 336
Db 181 caacttgagtgcccaagtgaatgataaagaacgatgaggttggtctcctcttcattatgcg 240
QY 337 gcttctgctggccgggatgagattgtaaaagccctctcgggaaaaaggtgctcaagtgaat 396
Db 241 gcttctgctggccgggatgagattgtaaaagccctctcgggaaaaaggtgctcaagtgaat 300
QY 397 gctgtcaatcaaaaaggctgtactcctctacattatgcagcttcgaaaaacacaggcatgag 456
Db 301 gctgtcaatcaaaaaggctgtactcctctacattatgcagcttcgaaaaacacaggcatgag 360
QY 457 atcgtgtcatgttactggaagcggggtaataatccagatgctaaagaccattatgagct 516
Db 361 atcgtgtcatgttactggaagcggggtaataatccagatgctaaagaccattatgagct 420
QY 517 acagcaatgcaccgggcagcagcgaagggttaacttgaagatgattcatatcctctgttac 576
Db 421 acagcaatgcaccgggcagcagcgaagggttaacttgaagatgattcatatcctctgttac 480
QY 577 tacaaagcatccacaacatccacagactgagggtgaacactcctctacacttagcctgt 636
Db 481 tacaaagcatccacaacatccacagactgagggtgaacactcctctacacttagcctgt 540
QY 637 gatgaggagagatggaagaagcaaaactgctggtgtcccaaggagcaagtattacatt 696
Db 541 gatgaggagagatggaagaagcaaaactgctggtgtcccaaggagcaagtattacatt 600
QY 697 gagaataaagaagaagaacacccctgcaagtggcccaagggtggtggtttaaatactc 756
Db 601 gagaataaagaagaagaacacccctgcaagtggcccaagggtggtggtttaaatactc 660
QY 757 aagagaatggtggaaggt 774
Db 661 aagagaatggtggaaggt 678

RESULT 4

AAAX35854

ID AAX35854 standard; DNA; 696 BP.

AC AAX35854;

14-JUL-1999 (first entry)

DNA encoding rat gankyrin protein.

Gankyrin; apoptosis induction; diagnosis; treatment; cancer;
hepatocellular carcinoma; oncogenesis mechanism; ss.

Rattus sp.

WO9918201-A1.

15-APR-1999.

02-OCT-1998; 98WO-JP04467.

03-OCT-1997; 97JP-0286214.

(FUJII/) FUJITA.

Fujita J;

WPI: 1999-277266/23.

P-PSDB; AAY02432.

Gankyrin polypeptides, useful for treatment and diagnosis of
cancers, e.g. hepatocellular carcinoma, and study of oncogenesis
mechanism

```
PS Claim 15; Page 75-76; 111pp; Japanese.
XX
CC The specification describes human, murine and rat gankyrin DNA and
CC polypeptide sequences. Gankyrin polypeptides inhibit tumorigenic
CC ability and apoptosis induction. The polypeptides and their antibodies
CC can be used in the diagnosis and treatment of cancers,
CC e.g. hepatocellular carcinoma, and study of oncogenesis mechanism.
CC The present sequence encodes rat gankyrin.
XX
SQ Sequence 696 BP; 211 A; 140 C; 183 G; 162 T; 0 other;

Query Match 76.0%; Score 593; DB 20; Length 696;
Best Local Similarity 91.0%; Pred. No. 3.7e-182;
Matches 626; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 97 atgaggggtgtgtctaaacctaatgtctgcaacctggtcctacagcggaagctggaa 156
Db 1 atgaggggtgtgtctaaacctaatgtctgcaacctggtcctacacgggaagctggat 60

Qy 157 gagttgaagagagatattctggtccgataataatccctggtctactagaactgaccaggacagc 216
Db 61 gagttgaaggaagcatttggctgataagctctgtgcccactagaactgatacaggacagc 120

Qy 217 agaactgattgactggcagctgctcagctggacatacagaaaattgtgaattttgtg 276
Db 121 agaacagcattgctgctggcagctgctcagctggacatacagaaaattgtgaattttgtg 180

Qy 277 caactggagtcaggtaataagagcagatgaggttggtctctctctctattgtg 336
Db 181 caactggagtcaggtaataagagcagatgaggttggtctctctctctattgtg 240

Qy 337 gcttctgctgcccggatgagattgtaaaagcccttctggaaaaggtgctcaagtgaat 396
Db 241 gcttctgctgcccggatgagattgtaaaagcccttctggaaaaggtgctcaagtgaat 300

Qy 397 gctgtcaatacaaaatggctgtactcccttaccattatcagcttcgaaaaacaggcatgag 456
Db 301 tctgtcaatacaaaagcgtgctgactcccttaccattatcagcttcgaaaaataggcatgag 360

Qy 457 atcgtgtcattgttactgggaagggcggtggttaataatccatgactaaggaccattaggcgt 516
Db 361 attctgttattgttactagaaggtggggtggttaataatccatgactaaggaccattaggcgt 420

Qy 517 acagcaatgcacccggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 576
Db 421 acagcaatgcacccggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 480

Qy 577 tacaaagcatccacaaacatccaaagacactgagggtaaacactcctctacacttagcctgt 636
Db 481 tacaaagcatccacaaacatccaaagacactgagggtaaacactcctctacacttagcctgt 540

Qy 637 gatgaggagagtggaagaagcaaaactgctggtgtcccaaggagcaagtattacatt 696
Db 541 gatgaggagagtggaagaagcaaaactgctggtgtcccaaggagcaagtattacatt 600

Qy 697 gagaataagaagaagaacacccctcgaagtggcgaaggtggcctgggttttaatactc 756
Db 601 gaaataaggaagaagaacacccctcgaagtggcgaaggtggcctgggttttaatactc 660

Qy 757 aagagaatggtggaagggttaa 777
Db 661 aagaagaatgcgaagaagttaa 681

RESULT 5
AA335853
ID AAX35853 standard; DNA; 696 BP.
XX
AC AAX35853;
XX
DT 14-JUL-1999 (first entry)
XX
```

```
DE DNA encoding mouse gankyrin protein.
XX
KW Gankyrin; apoptosis induction; diagnosis; treatment; cancer;
KW hepatocellular carcinoma; oncogenesis mechanism; ss.
XX
OS Mus sp.
XX
PN W09918201-A1.
XX
PD 15-APR-1999.
XX
PF 02-OCT-1998; 98WO-JP04467.
XX
PR 03-OCT-1997; 97JP-0286214.
XX
PA (FUJI/) FUJITA.
XX
PI Fujita J;
XX
DR WPI; 1999-277266/23.
DR P-PSDB; AAY02431.
XX
XX Gankyrin polypeptides, useful for treatment and diagnosis of
PT cancers, e.g. hepatocellular carcinoma, and study of oncogenesis
PT mechanism
XX
PS Claim 10; Page 71-73; 111pp; Japanese.
XX
CC The specification describes human, murine and rat gankyrin DNA and
CC polypeptide sequences. Gankyrin polypeptides inhibit tumorigenic
CC ability and apoptosis induction. The polypeptides and their antibodies
CC can be used in the diagnosis and treatment of cancers,
CC e.g. hepatocellular carcinoma, and study of oncogenesis mechanism.
CC The present sequence encodes mouse gankyrin.
XX
SQ Sequence 696 BP; 205 A; 142 C; 180 G; 168 T; 1 other;
```

```
Query Match 74.9%; Score 584; DB 20; Length 696;
Best Local Similarity 91.0%; Pred. No. 3e-179;
Matches 620; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 97 atgaggggtgtgtctaaacctaatgtctgcaacctggtcctacagcggaagctggaa 156
Db 1 atgaggggtgtgtctaaacctaatgtctgcaacctggtcctacagcggaagctggat 60

Qy 157 gagttgaagagagatattctggtccgataataatccctggtctactagaactgaccaggacagc 216
Db 61 gagttgaagagagcagcatttggctgataataatcctggtactagaactgacaggacagc 120

Qy 217 agaactgattgactggcagctgctcagctggacatacagaaaattgtgaattttgtg 276
Db 121 agaacagcattgctgctggcagctgctcagctggacatacagaaaattgtgaattttgtg 180

Qy 277 caactggagtcaggtaataagagcagatgaggttggtctctctctctattgtg 336
Db 181 caactggagtcaggtaataagagcagatgaggttggtctctctctctattgtg 240

Qy 337 gcttctgctgcccggatgagattgtaaaagcccttctggaaaaggtgctcaagtgaat 396
Db 241 gcttctgctgcccggatgagattgtaaaagcccttctggaaaaggtgctcaagtgaat 300

Qy 397 gctgtcaatacaaaatggctgtactcccttaccattatcagcttcgaaaaacaggcatgag 456
Db 301 tctgtcaatacaaaagcgtgctgactcccttaccattatcagcttcgaaaaataggcatgag 360

Qy 457 atcgtgtcattgttactgggaagggcggtggttaataatccatgactaaggaccattaggcgt 516
Db 361 attctgttattgttactagaaggtggggtggttaataatccatgactaaggaccattaggcgt 420

Qy 517 acagcaatgcacccggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 576
Db 421 acagcaatgcacccggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 480
```


XX PS Claim 1; Column 123-130; 84pp; English.
XX CC AAV62996-V63027 encode ankyrin proteins isolated from the helminth
XX CC parasites *Dirofilaria immitis* and *Brugia malayi*. The nucleic acids and
XX CC recombinant products are useful for the recombinant production of the
XX CC ankyrin polypeptides. These proteins can then be used as vaccines
XX CC against parasitic helminth, e.g. *D. immitis* or *B. malayi*. They can also
XX CC be used for therapy after infection, and to raise antibodies, also for
XX CC use in therapeutics, as passive immunogens, or as therapeutics against
XX CC helminths on conjugation to cytotoxic agents. The nucleic acids contained
XX CC in viruses, may also be used as viral vaccines, and the nucleic acids
XX CC themselves or in vectors may be used as genetic vaccines.
SQ Sequence 5235 BP; 1574 A; 1106 C; 1219 G; 1335 T; 1 other;

Query Match 12.0%; Score 93.4; DB 19; Length 5235;
Best Local Similarity 48.1%; Pred. No. 1.3e-19;
Matches 265; Conservative 0; Mismatches 286; Indels 0; Gaps 0;
QY 182 ataaatccctgctactagaactgaccagacagacagaaactgactgctggcagatgct 241
DB 1403 atggagacaggtggatgctgctgctggaactacaaactccactgacattgcatcac 1462
QY 242 cagctggacacacagaattgtgaattttgtgcaacttgagtgccagtgaaatgata 301
DB 1463 gtctggtaaccgacatcgctattgtgtgagtgatgacacaaatgctg 1522
QY 302 aagacgatcaggtgtgctctctctcatattgctgcttctgctgcccggatgagattg 361
DB 1523 ccacaagatctttatactctctctcatattgctgccaagagggcgaaggaagtgg 1582
QY 362 taaaagccctctgggaaagtgctcaagtgaatgctgcaatacaaatgctgtactc 421
DB 1583 cagcaatattggtgatcatggaacgacagacactgctcacgaaaggggtttacgc 1642
QY 422 ccttactattcagcttcgaaacagcagcatgagatcgctgctatgtttactggaagcg 481
DB 1643 cgttgctattgctgctgaagtatggcaatttccggctgcgaaatcattgctagaacgag 1702
QY 482 gggcctaataccagatgctgaaggaccattatgagctacagcaatgacccgggcagcgca 541
DB 1703 gaacacgggttgacattgaggaagaatcaggttaacacctctgcatgtagcgcaatt 1762
QY 542 agggtaacttgagatgattcatatctctctgtactacaagcatccacaacatccaag 601
DB 1763 acaataacgacagaagtgatgattgttacttctagaaaatggtctctgacatgccgctg 1822
QY 602 acactgagggtaacactctctacacttagcctgtgtagagagagagtggaagacaa 661
DB 1823 ccaagaatgggtacactctctacatattgctgcgcgaagaagatcagatcggtattgcta 1882
QY 662 aactgctggtgtcccaaggagcaagattttactatgagataaagaagaaaagacacccc 721
DB 1883 gcactctctctattataaggcaaatgcgaatgctgaaagcaagctggtttacacacc 1942
QY 722 tgaagtggtgcc 732
DB 1943 ttcatttgcc 1953

RESULT 9
AAV63023/c
ID AAV63023 standard; cDNA; 5235 BP.
XX
AC AAV63023;
XX
DT 15-JAN-1999 (first entry)
XX
DE D. immitis ankyrin cDNA complementary to nDIAnk5235.
XX
KW Ankyrin; helminth; parasite; vaccine; therapy; infection;

KW passive immunogen; cytotoxic agent; ss.
XX OS *Dirofilaria immitis*.
XX PN US5827692-A.
XX PD 27-OCT-1998.
XX PF 24-APR-1997; 97US-0847429.
XX PR 24-APR-1997; 97US-0847429.
XX PA (HESK-) HESKA CORP.
XX PI Blehm ES, Tang L;
XX DR WPI; 1998-593992/50.
XX Nucleic acids encoding ankyrins from helminth parasites - useful for
XX PT recombinant production of the proteins for use as vaccines and
XX PT treatments against helminth infection
XX PS Claim 1; Column 129-136; 84pp; English.
XX CC AAV62996-V63027 encode ankyrin proteins isolated from the helminth
XX CC parasites *Dirofilaria immitis* and *Brugia malayi*. The nucleic acids and
XX CC recombinant products are useful for the recombinant production of the
XX CC ankyrin polypeptides. These proteins can then be used as vaccines
XX CC against parasitic helminth, e.g. *D. immitis* or *B. malayi*. They can also
XX CC be used for therapy after infection, and to raise antibodies, also for
XX CC use in therapeutics, as passive immunogens, or as therapeutics against
XX CC helminths on conjugation to cytotoxic agents. The nucleic acids contained
XX CC in viruses, may also be used as viral vaccines, and the nucleic acids
XX CC themselves or in vectors may be used as genetic vaccines.
SQ Sequence 5235 BP; 1335 A; 1219 C; 1106 G; 1574 T; 1 other;

Query Match 12.0%; Score 93.4; DB 19; Length 5235;
Best Local Similarity 48.1%; Pred. No. 1.3e-19;
Matches 265; Conservative 0; Mismatches 286; Indels 0; Gaps 0;
QY 182 ataaatccctgctactagaactgaccagacagacagaaactgactgctggcagatgct 241
DB 3833 ATGGAGCAGAGTGGATGCTGCTGCTGGAACACTACAACACTCCACTGCACATTGCATCAC 3774
QY 242 cagctggacacacagaattgtgaattttgtgcaacttgagtgccagtgaaatgata 301
DB 3773 GTCTTGGTAATACCGACATCGTCAATTTGTGTGTCAGGCTAATGCATCACCATAATGCTG 3714
QY 302 aagacgatcaggtgtgctctctctcatattgctgcttctgctgcccggatgagattg 361
DB 3713 CCACAAGAGATCTTTATATCTCTCTTCAATTTGCTGCCAAGGAGGGCAGAGAAAGTGG 3654
QY 362 taaaagccctctgggaaagtgctcaagtgaatgctgcaatacaaatggcgtgactc 421
DB 3653 CAGCATATTTGATGGATCATGGACCGACAAGACACTGCTCAGCAAAAAGGGTTTACGC 3594
QY 422 ccttactattgacgcttcgaaaacagcagcatgagatcgctgctatgttactggaagcg 481
DB 3593 CGTTGCATTTAGCTGCTAAGTATGGCAATTTGCCGGTCGCGAAATCATTTGCTAGAACGAG 3534
QY 482 gggcctaataccagatgctgaaggaccattatgagctacagcaatgacccgggcagcgca 541
DB 3533 GAACACCCGGTTGACATTTGAAGGCAAGAAATCAGGTAAACACTCTGTCATGTAGCGGCACATT 3474
QY 542 agggtaacttgagatgattcatatctctgtactacaagcatccacaacatccaag 601
DB 3473 ACAATAACACAGGTAGCATTTGTTACTTCTAGAAAATGGTGTCTGCACATGCCGCTG 3414
QY 602 acactgagggtaacactctctacacttagcctgtgtagagagagagtggaagacaa 661
DB 3413 CCAAGAATGGGTACACTCCTTTTACATATTGCGCGAAGAAGAATCAGATGATGATATTGCTA 3354


```

XX The present sequence encodes a full length Dirofilaria immitis ankyrin
CC protein. The ankyrin protein, or anti-ankyrin antibodies, may be used to
CC protect an animal from disease caused by a parasitic helminth,
CC especially where the disease is heartworm disease, elephantiasis or
CC hydrocele.
XX
SQ Sequence 5503 BP; 1662 A; 1152 C; 1259 G; 1429 T; 1 other;

Query Match      12.0%; Score 93.4; DB 19; Length 5503;
Best Local Similarity 48.1%; Pred. No. 1.3e-19;
Matches 265; Conservative 0; Mismatches 286; Indels 0; Gaps 0;

QY 182 ataaatccctggctactagaaactgacacagacagagaactgcattgcaactggcgatgct 241
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1453 atggagcacaggtggatgctgctgctgctgctgctgctgctgctgctgctgctgctc 1512
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 242 cagctggacatacagaaattgtgaattttgttgcacacttgagtgccagtgaaatgata 301
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1513 gtcttggaataccgacatgctattttgttgcagtgctaatgcatcaccaaatgctg 1572
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 302 aagacgatgcaggttgctctctctctctctctctctctctctctctctctctctctc 361
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1573 ccacaagagatctttatactctctctctctctctctctctctctctctctctctctc 1632
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 362 taaaagccctctggaaaaggtgctcaagtgaatgctgtcaatcaaaatggctgtactc 421
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1633 cagcaatattgattgattgattgattgattgattgattgattgattgattgattgatt 1692
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 422 ccttactattgagatgattcattctctctctctctctctctctctctctctctctctc 481
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1693 cgttgatttagctgctgaattgctgaattgctgaattgctgaattgctgaattgctga 1752
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 482 gggctaataccagatgctgaaggaccattatgagctcacagcaatgcaccggcgagacca 541
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1753 gaacacgggttgacattggaaggcaagaatacagtaacacctctgcatgtagcgacatt 1812
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 542 agggtaacttgagatgattcattctctctctctctctctctctctctctctctctctc 601
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1813 acataacgacaaaggtagcattgttactctctctctctctctctctctctctctctc 1872
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 602 aactgaggttaaacctctctctctctctctctctctctctctctctctctctctctc 661
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1873 ccaagaatgggtacactcctctctctctctctctctctctctctctctctctctctc 1932
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 662 aactgctgggtgcccaaggacgaagtatttacctgagaaataaagaagaagaacacccc 721
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1933 gcactctctctctctctctctctctctctctctctctctctctctctctctctctc 1992
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 722 tgcaagtggcc 732
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1993 ttcatttgcc 2003
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
AAV63020
ID AAV63020 standard; cDNA; 5503 BP.
XX
AC AAV63020;
XX
XX 15-JAN-1999 (first entry)
DT
DE
DE D. immitis ankyrin nDiAnk5503 cDNA.
XX
KW Ankyrin; helminth; parasite; vaccine; infection;
KW passive immunogen; cytotoxic agent; ss.
XX
XX Dirofilaria immitis.
OS
XX Key Location/Qualifiers
FH 51..5288
FT CDS
FT /*tag= a

```

```

FT XX /product= "nDiAnk5503"
PN XX US5827692-A.
XX XX 27-OCT-1998.
XX XX 24-APR-1997; 97US-0847429.
XX XX 24-APR-1997; 97US-0847429.
XX XX (HESK-) HESKA CORP.
XX XX Blehm ES, Tang L;
XX XX WPI: 1998-593992/50.
XX XX P-PSDB; AAW76776.
PT Nucleic acids encoding ankyrins from helminth parasites - useful for
PT recombinant production of the proteins for use as vaccines and
PT treatments against helminth infection
XX PT
XX PS Claim 1; Column 95-108; 84pp; English.
XX XX
XX CC AAV62996-V63027 encode ankyrin proteins isolated from the helminth
XX CC parasites Dirofilaria immitis and Brugia malayi. The nucleic acids and
XX CC recombinant products are useful for the recombinant production of the
XX CC ankyrin polypeptides. These proteins can then be used as vaccines
XX CC against parasitic helminth, e.g. D. immitis or B. malayi. They can also
XX CC be used for therapy after infection, and to raise antibodies against
XX CC use in therapeutics, as passive immunogens, or as therapeutics against
XX CC helminths on conjugation to cytotoxic agents. The nucleic acids contained
XX CC in viruses, may also be used as viral vaccines, and the nucleic acids
XX CC themselves or in vectors may be used as genetic vaccines.
XX SQ Sequence 5503 BP; 1662 A; 1152 C; 1259 G; 1429 T; 1 other;

Query Match      12.0%; Score 93.4; DB 19; Length 5503;
Best Local Similarity 48.1%; Pred. No. 1.3e-19;
Matches 265; Conservative 0; Mismatches 286; Indels 0; Gaps 0;

QY 182 ataaatccctggctactagaaactgacacagacagagaactgcattgcaactggcgatgct 241
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1453 atggagcacaggtggatgctgctgctgctgctgctgctgctgctgctgctgctc 1512
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 242 cagctggacatacagaaattgtgaattttgttgcacacttgagtgccagtgaaatgata 301
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1513 gtcttggaataccgacatgctattttgttgcagtgctaatgcatcaccaaatgctg 1572
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 302 aagacgatgcaggttgctctctctctctctctctctctctctctctctctctctctc 361
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1573 ccacaagagatctttatactctctctctctctctctctctctctctctctctctctc 1632
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 362 taaaagccctctggaaaaggtgctcaagtgaatgctgtcaatcaaaatggctgtactc 421
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1633 cagcaatattgattgattgattgattgattgattgattgattgattgattgattgatt 1692
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 422 ccttactattgagatgattcattctctctctctctctctctctctctctctctctctc 481
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1693 cgttgatttagctgctgaattgctgaattgctgaattgctgaattgctgaattgctga 1752
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 482 gggctaataccagatgctgaaggaccattatgagctcacagcaatgcaccggcgagacca 541
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1753 gaacacgggttgacattggaaggcaagaatacagtaacacctctgcatgtagcgacatt 1812
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 542 agggtaacttgagatgattcattctctctctctctctctctctctctctctctctctc 601
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1813 acataacgacaaaggtagcattgttactctctctctctctctctctctctctctctc 1872
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 602 aactgaggttaaacctctctctctctctctctctctctctctctctctctctctctc 661
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1873 ccaagaatgggtacactcctctctctctctctctctctctctctctctctctctctc 1932
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 662 aactgctgggtgcccaaggacgaagtatttacctgagaaataaagaagaagaacacccc 721
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1933 gcactctctctctctctctctctctctctctctctctctctctctctctctctctc 1992
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 722 tgcaagtggcc 732
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1993 ttcatttgcc 2003
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
AAV63020
ID AAV63020 standard; cDNA; 5503 BP.
XX
AC AAV63020;
XX
XX 15-JAN-1999 (first entry)
DT
DE
DE D. immitis ankyrin nDiAnk5503 cDNA.
XX
KW Ankyrin; helminth; parasite; vaccine; infection;
KW passive immunogen; cytotoxic agent; ss.
XX
XX Dirofilaria immitis.
OS
XX Key Location/Qualifiers
FH 51..5288
FT CDS
FT /*tag= a

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2001, 07:46:08 ; Search time 80.36 Seconds
(without alignments)
1837.515 Million cell updates/sec

Title: US-09-509-775-1

Perfect score: 780

Sequence: 1 tggtagagctctaacgctg.....gaatggggaagttaaca 780

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_NA:*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*

2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*

3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*

4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*

5: /cgn2_6/ptodata/2/ina/6C_COMB.seq:*

6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	93.4	12.0	1056	3	US-09-065-474-138
2	93.4	12.0	1056	3	US-09-065-474-140
3	93.4	12.0	5235	1	US-09-031-485-35
4	93.4	12.0	5235	1	US-09-031-485-36
5	93.4	12.0	5235	1	US-08-847-429A-35
6	93.4	12.0	5235	1	US-08-847-429A-36
7	93.4	12.0	5235	3	US-09-065-474-35
8	93.4	12.0	5235	3	US-09-065-474-36
9	93.4	12.0	5503	1	US-09-031-485-32
10	93.4	12.0	5503	1	US-09-031-485-34
11	93.4	12.0	5503	1	US-08-847-429A-32
12	93.4	12.0	5503	1	US-08-847-429A-34
13	93.4	12.0	5503	3	US-09-065-474-32
14	93.4	12.0	5503	3	US-09-065-474-34
15	88.4	11.3	3454	4	US-09-082-059-1
16	86.6	11.1	911	1	US-09-031-485-22
17	86.6	11.1	911	1	US-09-031-485-24
18	86.6	11.1	911	1	US-08-847-429A-22
19	86.6	11.1	911	1	US-08-847-429A-24
20	86.6	11.1	911	3	US-09-065-474-22
21	86.6	11.1	911	3	US-09-065-474-24
22	85.6	11.0	909	1	US-09-031-485-25
23	85.6	11.0	909	1	US-09-031-485-26
24	85.6	11.0	909	1	US-08-847-429A-25
25	85.6	11.0	909	1	US-08-847-429A-26
26	85.6	11.0	909	3	US-09-065-474-25
27	85.6	11.0	909	3	US-09-065-474-26

28	78.6	10.1	908	1	US-09-031-485-37	Sequence 37, Appl
c 29	78.6	10.1	908	1	US-09-031-485-39	Sequence 39, Appl
30	78.6	10.1	908	1	US-08-847-429A-37	Sequence 37, Appl
c 31	78.6	10.1	908	1	US-08-847-429A-39	Sequence 39, Appl
32	78.6	10.1	908	3	US-09-065-474-37	Sequence 37, Appl
c 33	78.6	10.1	908	3	US-09-065-474-39	Sequence 39, Appl
34	77.6	9.9	906	1	US-09-031-485-40	Sequence 40, Appl
c 35	77.6	9.9	906	1	US-09-031-485-41	Sequence 41, Appl
36	77.6	9.9	906	1	US-08-847-429A-40	Sequence 40, Appl
c 37	77.6	9.9	906	1	US-08-847-429A-41	Sequence 41, Appl
38	77.6	9.9	906	3	US-09-065-474-40	Sequence 40, Appl
c 39	77.6	9.9	906	3	US-09-065-474-41	Sequence 41, Appl
40	74.4	9.5	573	1	US-09-031-485-19	Sequence 19, Appl
c 41	74.4	9.5	573	1	US-09-031-485-21	Sequence 21, Appl
42	74.4	9.5	573	1	US-08-847-429A-19	Sequence 19, Appl
c 43	74.4	9.5	573	1	US-08-847-429A-21	Sequence 21, Appl
44	74.4	9.5	573	3	US-09-065-474-19	Sequence 19, Appl
c 45	74.4	9.5	573	3	US-09-065-474-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-09-065-474-138
; Sequence 138, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 138:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1056 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1056
US-09-065-474-138

Query Match 12.0%; Score 93.4; DB 3; Length 1056;
Best Local Similarity 48.1%; Pred. No. 8.3e-21;

Matches 265; Conservative 0; Mismatches 286; Indels 0; Gaps 0;

QY 182 ataaatccctggctactagaaactgaccagacagacgaactgcaactggcgatgct 241
II III III III III III III III III III III III III III III III
Db 98 ATGAGCACAGGTGGATGCTGCTCGTGAACACTACAACTCCACTGCACATTCATCAC 157
II III III III III III III III III III III III III III III III
QY 242 cagctgacatacagaattgtgaattttgttgcaacttgagtgccagtgaaatgata 301
II III III III III III III III III III III III III III III III
Db 158 GTCTTGGTAATACCGACATGCTATTTGTTGCTGAGGCTAATGCATCACCAGTCTG 217
II III III III III III III III III III III III III III III III
QY 302 aagcagtgaggtgtgctcccttcctcatattgcggtcttctgctgcccggatgagattg 361
II III III III III III III III III III III III III III III III
Db 218 CCACAAGAGATCTTATACCTCTTCTCATATTGCTGCAAGAGGGGCAAGAGATGG 277
II III III III III III III III III III III III III III III III
QY 362 taaagcccttctgggaaagggtgctcaagtgaatgctgcaatcaaaatggctgtactc 421
II III III III III III III III III III III III III III III III
Db 278 CAGCAATATTGATGATCATGGAACCCGACAGACACTGCTCACGAAAAGGGTTTACGC 337
II III III III III III III III III III III III III III III III
QY 422 ccttacattatgcagcttcgaataaacaggcatgagatcgctgtcatgttactggaagcg 481
II III III III III III III III III III III III III III III III
Db 338 CGTTGCAATTTAGCTGCTAAGTATGGCAATTTGCCGGTCCGGGAAATCATTTGCTAGACGAG 397
II III III III III III III III III III III III III III III III
QY 482 gggctaataccagatgctaagaccattatgagctacagcaatgcacccggcagagcca 541
II III III III III III III III III III III III III III III III
Db 398 GAACACCGTTGACATTAAGGCAAGAAATCAGGTAACACCTCTGCATGTAGCGGCACATT 457
II III III III III III III III III III III III III III III III
QY 542 agggtaactggaagatgattcatatcctctgtactacaagcatccacaaacatccaag 601
II III III III III III III III III III III III III III III III
Db 458 ACAATAACGACAAGTAGCATTTACTTCTAGAAAATGGTCTTCTGCATCCTCCGCTG 517
II III III III III III III III III III III III III III III III
QY 602 acactgagggtaacactcctctacactgctgtgtatgagggagagtggaagaagcaa 661
II III III III III III III III III III III III III III III III
Db 518 CCAAGAATGGGTACACTCCTTTACATATTGCCGGAAGAAATCAGATGGATATTGCTA 577
II III III III III III III III III III III III III III III III
QY 662 aactgctgggtcccaaggagcaagtattacattgagaataaagaagaagacacccc 721
II III III III III III III III III III III III III III III III
Db 578 GCATCTCCTTCATTATTAAGCAAAATCGGAATGCTGAAAGCAAAAGCTGGCTTTACACCAC 637
II III III III III III III III III III III III III III III III
QY 722 tgcaagtggcc 732
II III III
Db 638 TTCATCTTGC 648
II III III

RESULT 2

US-09-065-474-140/c
; Sequence 140, Application US/09065474
; Patent No. 6063599

GENERAL INFORMATION:

; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171

CORRESPONDENCE ADDRESS:

; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Hesk Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington

; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 140:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1056 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-065-474-140

Query Match 12.0%; Score 93.4; DB 3; Length 1056;

Best Local Similarity 48.1%; Pred. No. 8.3e-21;

Matches 265; Conservative 0; Mismatches 286; Indels 0; Gaps 0;

QY 182 ataaatccctggctactagaaactgaccagacagacgaactgcaactggcgatgct 241
II III III III III III III III III III III III III III III III
Db 959 ATGAGCACAGGTGGATGCTGCTCGTGAACACTACAACTCCACTGCACATTCATCAC 900
II III III III III III III III III III III III III III III III
QY 242 cagctgacatacagaattgtgaattttgttgcaacttgagtgccagtgaaatgata 301
II III III III III III III III III III III III III III III III
Db 899 GTCTTGGTAATACCGACATGCTCATTTTGTGCTGCAAGGCTAATGCATCACCAGTCTG 840
II III III III III III III III III III III III III III III III
QY 302 aagcagtgaggtgtgctcccttcctcatattgcggtcttctgctgcccggatgagattg 361
II III III III III III III III III III III III III III III III
Db 839 CCACAAGAGATCTTATACCTCTTCTCATATTGCTGCCAAGAGGGGCAAGAGGATGG 780
II III III III III III III III III III III III III III III III
QY 362 taaagcccttctgggaaagggtgctcaagtgaatgctgcaatcaaaatggctgtactc 421
II III III III III III III III III III III III III III III III
Db 779 CAGCAATATTGATGATCATGGAACCCGACAGACACTGCTCACGAAAAGGGTTTACGC 720
II III III III III III III III III III III III III III III III
QY 422 ccttacattatgcagcttcgaataaacaggcatgagatcgctgtcatgttactggaagcg 481
II III III III III III III III III III III III III III III III
Db 719 CGTTGCAATTTAGCTGCTAAGTATGGCAATTTGCCGGTCCGGGAAATCATTTGCTAGACGAG 660
II III III III III III III III III III III III III III III III
QY 482 gggctaataccagatgctaagaccattatgagctacagcaatgcacccggcagagcca 541
II III III III III III III III III III III III III III III III
Db 659 GAACACCGTTGACATTAAGGCAAGAAATCAGGTAACACCTCTGCATGTAGCGGCACATT 600
II III III III III III III III III III III III III III III III
QY 542 agggtaactggaagatgattcatatcctctgtactacaagcatccacaaacatccaag 601
II III III III III III III III III III III III III III III III
Db 599 ACAATAACGACAAGGTAGCATTTACTTCTAGAAAATGGTCTTCTGCACATGCCGCTG 540
II III III III III III III III III III III III III III III III
QY 602 acactgagggtaacactcctctacactgctgtgtatgagggagagtggaagaagcaa 661
II III III III III III III III III III III III III III III III
Db 539 CCAAGAATGGGTACACTCCTTTACATATTGCCGGAAGAAATCAGATGGATATTGCTA 480
II III III III III III III III III III III III III III III III
QY 662 aactgctgggtcccaaggagcaagtattacattgagaataaagaagaagacacccc 721
II III III III III III III III III III III III III III III III
Db 479 GCATCTCCTTCATTATTAAGCAAAATCGGAATGCTGAAAGCAAAAGCTGGCTTTACACCAC 420
II III III III III III III III III III III III III III III III
QY 722 tgcaagtggcc 732
II III III
Db 419 TTCATCTTGC 409
II III III

RESULT 3

US-09-031-485-35
; Sequence 35, Application US/09031485
; Patent No. 5824306

GENERAL INFORMATION:

; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.

OM of: US-09-509-775-1 to: PIR_68:* out_format : pfs

Date: Aug 13, 2001 8:34 AM

About: Results were produced by the GenCore software, version 4.5, Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

```
-MODEL=framet_n2p_model -DEVS=xlp
-O=/csgn2_1/USPTO-specol/USO9509775/runat_13082001_074528_13056/app_query.fasta_1.848
-DB=PIR_58 QFMT=fastan -SUFFIX=tpi GAPOP=12.000 GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOPEXT=0.000 -GGAPOP=4.500
-GGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-DEGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DLOCAL=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODEL=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=2000000000 -USER=USO9509775_@cgn1_1_64 -NCPU=3
-LOGLOG -NO_XLPXY -WAIT -THREADS=51
```

Search information block:

Query: US-09-509-775-1

Query length: 780

Database: PIR_68:*

Database sequences: 219241

Database length: 76174552

Search time (sec): 92.960000

score_list:	Strd	Orig	zScore	EScore	Len	Documentation
p1r2:084448	+	338.00	560.20	1.3e-23	247	probable ankryn [imported] - mouse
p1r2:084449	+	312.50	500.61	3.8e-21	1765	ankryn 3, splice form 2 - mouse
p1r2:742715	+	312.50	499.82	3.9e-21	1940	ankryn 3, splice form 3 - mouse
p1r2:742716	+	312.50	499.80	3.9e-21	1943	ankryn 3, splice form 1 - mouse
p1r2:742717	+	312.50	499.73	3.9e-21	1961	ankryn 3, splice form 4 - mouse
p1r2:750964	+	310.00	513.29	5.6e-21	237	related to 26s proteasome subunit
p1r2:A55757	+	310.00	486.75	7.1e-21	4377	ankryn 3, long splice form - mouse
p1r2:T33631	+	305.00	496.54	1.8e-20	636	hypothetical protein F40G9.1 - mouse
p1r2:537431	+	294.00	462.66	2.4e-19	3924	ankryn 2, neuronal long splice
p1r2:557697	+	291.00	481.54	3.4e-19	228	hypothetical protein YGR232W - mouse
p1r2:B35049	+	290.00	462.21	5.0e-19	1856	ankryn 1, erythrocyte splice
p1r2:A35049	+	290.00	462.10	5.0e-19	1880	ankryn 1, erythrocyte splice
p1r1:SUHUK	+	290.00	462.10	5.0e-19	1881	ankryn 1, erythrocyte splice
p1r2:A57282	+	288.50	460.00	6.9e-19	1786	ankryn-related protein unc-44
p1r2:15346	+	288.50	459.87	7.0e-19	1815	elegans ankryn-related unc-44
p1r2:15344	+	288.50	459.63	7.0e-19	1867	ankryn-related unc-44 - Caenorhabditis
p1r2:15347	+	288.50	458.89	7.0e-19	2039	ankryn-related unc-44 - Caenorhabditis
p1r2:537771	+	287.50	458.03	8.7e-19	1848	ankryn, erythrocyte - mouse
p1r2:149502	+	287.50	457.97	8.7e-19	1862	ankryn - mouse
p1r2:133940	+	276.50	440.95	9.2e-18	1549	ankryn - fruit fly (Drosophila)
p1r2:T39032	+	273.50	451.79	1.5e-17	234	hypothetical ankryn-like protein
p1r2:Q0197	+	270.50	436.22	3.2e-17	815	myosin-light-chain-phosphatase
p1r2:568418	+	269.50	436.34	3.9e-17	658	protein phosphatase 1M chain ML
p1r2:A55142	+	268.50	431.09	5.0e-17	1004	myosin-light-chain-phosphatase
p1r2:T30255	+	263.00	421.34	1.7e-16	1062	inversin - mouse
p1r2:14151	+	246.00	409.52	7.6e-16	1062	Inv protein - mouse
p1r2:746445	+	238.50	405.14	3.6e-15	397	hypothetical protein DKF2p434B2
p1r2:S11527	+	247.50	392.85	4.9e-15	1401	alpha-latrotoxin precursor - human
p1r1:173275	+	247.00	391.87	5.4e-15	1423	death-associated protein kinase
p1r2:530355	+	243.50	386.03	1.2e-14	1411	alpha-latroinsectotoxin precursor
p1r2:18184	+	239.50	390.59	2.5e-14	368	ankryn repeat protein A682L - mouse
p1r2:A57291	+	239.00	390.95	2.7e-14	319	cytokine inducible nuclear protein
p1r2:746507	+	239.00	386.26	2.9e-14	557	hypothetical protein DKF2p58M2
p1r2:742691	+	237.50	380.78	4.1e-14	791	hypothetical protein DKF2p434D2
p1r2:125053	+	237.50	380.16	4.1e-14	851	hypothetical protein DKF2p434A1
p1r2:101729	+	236.50	384.04	4.8e-14	439	ankryn-repeat protein - Arabidopsis
p1r2:743458	+	234.50	373.48	7.9e-14	1031	hypothetical protein DKF2p434H
p1r1:7171274	+	233.00	371.78	1.1e-13	934	probable ankryn - syphilis sp.
p1r2:721884	+	232.00	366.70	1.4e-13	1398	hypothetical protein F36H1.2 - mouse
p1r2:191952	+	229.00	363.01	2.6e-13	1188	hypothetical protein C29B6.2 - mouse
p1r2:191906	+	228.00	362.64	3.2e-13	1016	ankryn related protein C06G3.1 - mouse

```

381 .....GCTGCTCAAGTGAATCTGCTGCTCAATCAAAATGGCTGTACT 420
583 rProLeuHisValAlaAlaHisTyrAspAsnGlnLysValAlaLeuLeuL 600
382 .....GCTGCTCAAGTGAATCTGCTGCTCAATCAAAATGGCTGTACT 420
600 euLeuAspGlnGlyAlaSerProHisAlaAlaAlaLysAsnGlyTyrThr 616
421 CCCTTACTATTGACGCTTCGAAACACAGCGCATGAGATCGCTGTCATGTT 470
617 ProLeuHisIleAlaAlaLysLysAsnGlnMetAspIleAlaThrSerLe 633
471 ACTGGGAAGCGGGCTAAATCCAGATCTCTAAGGACCATTTATGAGGCTACAG 520
633 uLeuGluTyrGlyAlaAspAlaAsnAlaValThrGlnGlnGlyIleAlas 650
521 CAATGCACCGGGCAGCAGCAAGGGTAACTTGAAGATGATTCAATTCCTT 570
650 erValHisLeuAlaGlnGluGlyHisValAspMetValSerLeuLeu 666
571 CTGTACTACAAGCATCCACAACATCCAGACACTGAGGGTAAACACTCC 620
667 LeuSerArgAsnAlaAsnValAsnLeuSerAsnLysSerGlyLeuThrPr 683
621 TCTACACTTACGCTGTATGATGAGGAGAGTGGGAAGAACAAACTCTGG 670
683 oLeuHisLeuAlaGlnGluAspArgValAsnValAlaGluValLeuV 700
671 TGTCCTCAAGGACGCAAGTATTTCATTTGAGAGTAAGAAGAAAAGACACC 720
700 aLAsnGlnGlyAlaHisValAspAlaGlnThrLysMetGlyTyrThrPro 716
721 CTGCAAGTG 729
717 LeuHisVal 719

seq_name: pir2:T42715
seq_documentation_block:
ankyrin 3, splice form 3 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change
C;Accession: T42715
R;Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.
J. Cell Biol. 130, 313-330, 1995
A;Title: Ank3 (epithelial ankyrin), a widely distributed new mem
the repeat domain.
A;Reference number: Z22237; MUID:95340633
A;Accession: T42715
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1940 <PEP>
A;Cross-references: EMBL:L40632; NID:g710548; PID:g710549; PIDID
A;Experimental source: strain C57BL/6J; kidney
C;Genetics:
A;Gene: Ank3
A;Map position: 10
A;Introns: 834/1
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing

alignment_scores:
Quality: 312.50 Length: 203
Ratio: 2.500 Gaps: 1
Percent Similarity: 61.576 Percent Identity: 35.468

alignment_block:
US-09-509-775-1 x T42715 ..
Align seg 1/1 to: T42715 from: 1 to: 1940

```

```

381 .....GCTGCTCAAGTGAATCTGCTGCTCAATCAAAATGGCTGTACT 420
583 rProLeuHisValAlaAlaHisTyrAspAsnGlnLysValAlaLeuLeuL 600
382 .....GCTGCTCAAGTGAATCTGCTGCTCAATCAAAATGGCTGTACT 420
600 euLeuAspGlnGlyAlaSerProHisAlaAlaAlaLysAsnGlyTyrThr 616
421 CCCTTACTATTGACGCTTCGAAACACAGCGCATGAGATCGCTGTCATGTT 470
617 ProLeuHisIleAlaAlaLysLysAsnGlnMetAspIleAlaThrSerLe 633
471 ACTGGGAAGCGGGCTAAATCCAGATCTCTAAGGACCATTTATGAGGCTACAG 520
633 uLeuGluTyrGlyAlaAspAlaAsnAlaValThrGlnGlnGlyIleAlas 650
521 CAATGCACCGGGCAGCAGCAAGGGTAACTTGAAGATGATTCAATTCCTT 570
650 erValHisLeuAlaGlnGluGlyHisValAspMetValSerLeuLeu 666
571 CTGTACTACAAGCATCCACAACATCCAGACACTGAGGGTAAACACTCC 620
667 LeuSerArgAsnAlaAsnValAsnLeuSerAsnLysSerGlyLeuThrPr 683
621 TCTACACTTACGCTGTATGATGAGCAGAGATGGGAAGAACAAACTCTGG 670
683 oLeuHisLeuAlaGlnGluAspArgValAsnValAlaGluValLeuV 700
671 TGTCCTCAAGCAGCAAGTATTATCATTTGAGAGTAAGAAGAAAAGACACC 720
700 aLAsnGlnGlyAlaHisValAspAlaGlnThrLysMetGlyTyrThrPro 716
721 CTGCAAGTG 729
717 LeuHisVal 719

seq_name: pir2:T42715
seq_documentation_block:
ankyrin 3, splice form 3 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_changed
C;Accession: T42715
R;Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.
J. Cell Biol. 130, 313-330, 1995
A;Title: Ank3 (epithelial ankyrin), a widely distributed new mem
the repeat domain.
A;Reference number: Z22237; MUID:95340633
A;Accession: T42715
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1940 <PEP>
A;Cross-references: EMBL:L40632; NID:g710548; PID:g710549; PIDN
A;Experimental source: strain C57BL/6J; kidney
C;Genetics:
A;Gene: Ank3
A;Map position: 10
A;Introns: 834/1
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing

alignment_scores:
Quality: 312.50 Length: 203
Ratio: 2.500 Gaps: 1
Percent Similarity: 61.576 Percent Identity: 35.468

alignment_block:
US-09-509-775-1 x T42715
Align seg 1/1 to: T42715 from: 1 to: 1940

```

```

517 ThrProLeuHisLeuAlaAlaArgGluGlyHisGluAspValAlaAlaPh 533
270 TTTTCTTCCAACCTGGAGTGCAGCTGAATGATAAAGACGATGCAGTTGGT 319
533 eLeuLeuAspHisGlyAlaSerLeuSerIleThrThrLysLysGlyPheT 550
320 CTCTCTTCATATTGCGCTTCTGCTGCGCGGATGAGATTGTAAGAAGCC 369
550 hrProLeuHisValAlaAlaLysTyrglyLysLeuGluValAlaSerLeu 566
370 CTCTCTGGGAAAA 381
567 LeuLeuGlnLysSerAlaSerProAspAlaAlaGlyLysSerGlyLeuTh 583
381 ..... 381
583 rProLeuHisValAlaAlaHisTyraSpAsnGlnLysValAlaLeuLeuL 600
382 .....GGTGTCAAGTGAATGCTGCTCAATCAAAATGGCTGTACT 420
600 euLeuAspGlnGlyAlaSerProHisAlaAlaLysAsnGlyTyThr 616
421 CCCTTACATTATGAGCTTCGAAAACAGCATGAGATCGCTGTCAATGTT 470
617 ProLeuHisIleAlaAlaLysLysAsnGlnMetAspIleAlaThrSerLe 633
471 ACTGGAAGCGGGCTTAATCCAGATGCTAAGGACCATATTAGCGCTACAG 520
633 uLeuGluTyrglyAlaAlaLysAlaAlaValThrArgGlnGlyIleAla 650
521 CAATGCACCGCGGAGCAGCAGCAAGGTAACCTGAAGATGATTTCATCTT 570
650 erValHisLeuAlaAlaGlnGlyHisValAspMetValSerLeuLeu 666
571 CTGTACTACAAAGCATCCACAAACATCCAAAGACACTGAGGGTAAACACTCC 620
667 LeuSerArgAsnAlaAsnValAsnLeuSerAsnLysSerGlyLeuThrPr 683
621 TCACACTTACCTGTGATGAGAGAGAGTGGGAAGCAAACTGCTGG 670
683 oLeuHisLeuAlaAlaGlnGluAspArgValAsnValAlaGluValLeuV 700
671 TGTCCTCCAGGAGCAAGTATTACATTTCAGATAAAGAAAGAAAGACACCC 720
700 alAsnGlnGlyAlaHisValAspAlaGlnThrLysMetGlyTyThrPro 716
721 CTGCAAGTG 729
717 LeuHisVal 719

```

seq_name: pir2:T42713

```

seq_documentation_block:
ankyrin 3, splice form 1 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C:Accession: T42713
R:Peters, L.I.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Vialamas, M.; Turtzo, J.
J. Cell Biol. 130, 313-330, 1995
A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene
the repeat domain.
A:Reference number: Z22237; MUID:95340633
A:Accession: T42713
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1943 <PET>
A:Cross-references: EMBL:L40632; NID:g710548; PID:g710550; PIDN:AAB01606.1
A:Experimental source: strain C57BL/6J; kidney
C:Genetics:
A:Gene: Ank3
A:Map position: 10
A:Introns: 855/1

```

C:Function:
A:Description: supposed to play an important role in the polarized distribution of ma
A>Note: major kidney ankyrin
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing

alignment_scores:
Quality: 312.50 Length: 203
Ratio: 2.500 Gaps: 1
Percent Similarity: 61.576 Percent Identity: 35.468
alignment_block:
US-09-509-775-1 x T42713 ..

Align seg 1/1 to: T42713 from: 1 to: 1943

```

220 ACTGCATTCAGCTGGGCATCTCAGCTGGACATACAGAAATTTGTAATT 269
517 ThrProLeuHisLeuAlaAlaArgGluGlyHisGluAspValAlaAlaPh 533
270 TTTTCTTCCAACCTGGAGTGCAGCTGAATGATAAAGACGATGCAGTTGGT 319
533 eLeuLeuAspHisGlyAlaSerLeuSerIleThrThrLysLysGlyPheT 550
320 CTCTCTTCATATTGCGCTTCTGCTGCGCGGATGAGATTGTAAGAAGCC 369
550 hrProLeuHisValAlaAlaLysTyrglyLysLeuGluValAlaSerLeu 566
370 CTCTCTGGGAAAA 381
567 LeuLeuGlnLysSerAlaSerProAspAlaAlaGlyLysSerGlyLeuTh 583
381 ..... 381
583 rProLeuHisValAlaAlaHisTyraSpAsnGlnLysValAlaLeuLeuL 600
382 .....GGTGTCAAGTGAATGCTGCTCAATCAAAATGGCTGTACT 420
600 euLeuAspGlnGlyAlaSerProHisAlaAlaLysAsnGlyTyThr 616
421 CCCTTACATTATGAGCTTCGAAAACAGCATGAGATCGCTGTCAATGTT 470
617 ProLeuHisIleAlaAlaLysLysAsnGlnMetAspIleAlaThrSerLe 633
471 ACTGGAAGCGGGCTTAATCCAGATGCTAAGGACCATATTAGCGCTACAG 520
633 uLeuGluTyrglyAlaAlaLysAlaAlaValThrArgGlnGlyIleAla 650
521 CAATGCACCGCGGAGCAGCAGCAAGGTAACCTGAAGATGATTTCATCTT 570
650 erValHisLeuAlaAlaGlnGlyHisValAspMetValSerLeuLeu 666
571 CTGTACTACAAAGCATCCACAAACATCCAAAGACACTGAGGGTAAACACTCC 620
667 LeuSerArgAsnAlaAsnValAsnLeuSerAsnLysSerGlyLeuThrPr 683
621 TCACACTTACCTGTGATGAGAGAGAGTGGGAAGCAAACTGCTGG 670
683 oLeuHisLeuAlaAlaGlnGluAspArgValAsnValAlaGluValLeuV 700
671 TGTCCTCCAGGAGCAAGTATTACATTTCAGATAAAGAAAGAAAGACACCC 720
700 alAsnGlnGlyAlaHisValAspAlaGlnThrLysMetGlyTyThrPro 716
721 CTGCAAGTG 729
717 LeuHisVal 719

```

seq_name: pir2:T42716

seq_documentation_block:
ankyrin 3, splice form 4 - mouse

C;Accession: I3V984
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu
submitted to the Protein Sequence Database, July 2000

A:Accession: T50984
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-237 <SCH>
A:cross-references: EMBL:AL389891; GSPDB:GN00116; NCSP:B7F18.30
A:Experimental source: BAC clone B7F18; strain OR74A
C:Genetics:
A:Gene: NCSP:B7F18.30
A:Map position: 6
A:Introns: 17/1; 25/3; 68/3

A;Map position: 6
A;Introns: 17/1; 25/3; 68/3

```

alignment_scores:
  Quality: 310.00      Length: 226
  Ratio: 2.279        Gaps: 4
  Percent Similarity: 60.177      Percent Identity: 34.513

alignment_block:
  US-09-509-775-1 x T50984      ..

Align seg 1/1 to: T50984 from: 1 to: 237

```

[illegible]

28 oLysLeuIaGlnArgIysAspAspGlyArgLeuProIleHisTrpA 45
236 CATGCTCAGCTGGACATACAGAAATGTTGAATTTTGTG...CAACTT 292

283 GGAGTGGCCAGTAGATGAAGAAGCATGCAGGTGCTCCTCTTCATAT
||| :::: |||||| | ||||| : : : :
62 GlyPheAspProAspValGIuaSPasMeIGlyThrPrroPheMetII 78

333 TCGGGCTTCTCT...GCCCGGAGAGATGTAAAGCCCTTCCTTCGGAA
78 eSerAlaSerValIysAspSerAspAlaIleAspLeuLeuSerA 95
380 AAGTGTCTCAAGTGAATGCTGTCAATCAAATAGCTGTACTCCCTTACAT 429

95 rgGlyAlaAspIleAsnGlnThrAsnHisGlnSerGlnThrAlaLeuHis 1111
430 TATGCAGCTTCGAAACACAGCATGCATGCCGTCTCATGTTACTTGAAGG 479

480 CGGGGCTAATCCAGT.....GCTAAGACCATTCAGG 514
 ::||| :::|||||

154 GAAGAGTTGAAGGACGATATTCTGGCGATAAATCCCTGGCTACTAGAAC 203
|||:::||||:::||||:::||||:::
545 GlusapValAlaAlaPheLeuLeuAspHisGlyAlaSerLeuSerIleTh 561
|||:::||||:::||||:::||||:::
204 TGACCAGGCACGCAAACTGCATTTCGCTGGCATGCTCAGCTGGACATA 253
|||:::||||:::||||:::||||:::
561 rThrLysylsilylphenylProLeuHisValAlaAlaLysTyrGlyLysL 578
|||:::||||:::||||:::||||:::
254 CAGAAATGTGTGAATTTTGTTCGAACCTGGAGCTGCCAGTGAAATGATAA 303
|||:::||||:::||||:::||||:::
578 euGluValAlaAsnLeuLeuGlnLysSerAlaSerProAspAlaAla 594
|||:::||||:::||||:::||||:::
304 GACGATCGAGTTGGTCTCTCTTCATATTGGCGGTCTCTGCTGCCGGGA 353
|||:::||||:::||||:::||||:::

#: 75-621/ДОМаш: анкугилл гербат номлогч <АНЗ3>

A;Status: preliminary; translated

Session: T33631

Session: T33631

1 MetSerAsnTyrProLeuHisGlnAlaCysMetGluAsnGluPhePhe 17
159 GTTGAAGAGAGATATTCTGGCGATAAATCCCTGCTACTAGAACTGACC 208
17 sValGlnGluLeuLeuHisSerLysProSerLeuLeuGlnLysAspG 34
209 AGGACAGAGAACTGCATGCTACTGGGATGCTCAGCTGGACATACAGAA 258
34 InAspGlyArgIleProLeuHisLysTrpSerValSerPheGlnAlaHisGlu 50
259 ATTGTTGAATTTTCTTCACTTGGAGTGGCCAGTGAATGA 299
51 lleThrSerPheLeuLeuLysMetGluAsnValAsnLeuAspAspTyr 67
300 TAAAGACATGTCAGGCTGCTCTCTCTCATATTGCGGCTTCTGCTGCC 349
67 rProAspSerGlyTrpThrProPheHisIleAlaCysSerValGlyAla 84
350 GGGATGAGATGTTAAAGCCCTTCTGGGAAAGGTGCTCAAGTG 393
84 snLeuGluValValLysSerLeuLysTrpSerValSerPheGlnAlaHisGlu 100
394 AATGCTGTCAATCAAAAGGCTGCTACTCCCTTACATATTGAGCTTCGAA 443
101 AsnLysIleThrAsnGlnGlyValThrCysLeuHisLeuAlaValGlyLy 117
444 AAACAGCATGATGATGCTGCTCATGTTACTGGAAGCGGGCTAATCCAG 493
117 sLysTrpPheGluValSerGlnPheLeuIleGluAsnGlyAlaSerVala 134
494 ATGTAAGGACCAATTATGAGCTACAGCAATGACCGGGCGAGCAGCCCAAG 543
134 rGleLysAspLysPheAsnGlnIleProLeuHisArgAlaAlaSerVal 150
544 GGTAACTTGAAGATGATTCATATCTCTCTGCTGCTGCTGCTGCTGCTG 590
151 GlySerLeuLysLeuIleGluLeuLysCysGlyLeuGlyLysSerAlaVala 167
591 AAACATCAAGACACTGAGGTAACTCTCTCTACACTTACCTGCTGATG 640
167 lAsnTrpGlnAspLysGlnGlyTrpThrProLeuPheHisAlaLeuAlaG 184
641 AGGAGAGTGGAGAGCAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 690
184 luGlyHisGlyAspAlaAlaValLeuLeuValGluLysTrpGlyAlaGlu 200
691 TACATTGAGAATAAAGAGAAAG 714
201 TyrAspLeuValAspAsnLysGlyAlaLys 210

seq_name: pir2:B35049

seq_documentation_block:

N:ankyrin 1, erythrocyte splice form 3 - human
N:alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
N:Contains: ankyrin 2.2, erythrocyte
C:Species: Homo sapiens (man)
C:Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 10-Jul-1998
R:Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.; K
proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A:title: cDNA sequence for human erythrocyte ankyrin.
A:Reference number: A35049; MUID:90175370
A:Accession: B35049
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1856 <LAW>
C:Genetics:
A:Gene: GDB:ANK1; ANK
A:Cross-references: GDB:118737; OMIM:182900
A:Map position: 8p11.2-8p11.2
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing

F;2-1856/Product: ankyrin 1, erythrocyte form 3 #status predicted <MAT>
F;2-1513,1676-1856/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>
F;44-76/Domain: ankyrin repeat homology <AN01>
F;77-109/Domain: ankyrin repeat homology <AN02>
F;110-142/Domain: ankyrin repeat homology <AN03>
F;143-171/Domain: ankyrin repeat homology <AN04>
F;172-204/Domain: ankyrin repeat homology <AN05>
F;205-237/Domain: ankyrin repeat homology <AN06>
F;238-270/Domain: ankyrin repeat homology <AN07>
F;271-303/Domain: ankyrin repeat homology <AN08>
F;304-336/Domain: ankyrin repeat homology <AN09>
F;337-369/Domain: ankyrin repeat homology <AN10>
F;370-402/Domain: ankyrin repeat homology <AN11>
F;403-435/Domain: ankyrin repeat homology <AN12>
F;436-468/Domain: ankyrin repeat homology <AN13>
F;469-501/Domain: ankyrin repeat homology <AN14>
F;502-534/Domain: ankyrin repeat homology <AN15>
F;535-567/Domain: ankyrin repeat homology <AN16>
F;568-600/Domain: ankyrin repeat homology <AN17>
F;601-633/Domain: ankyrin repeat homology <AN18>
F;634-666/Domain: ankyrin repeat homology <AN19>
F;667-699/Domain: ankyrin repeat homology <AN20>
F;700-732/Domain: ankyrin repeat homology <AN21>
F;733-765/Domain: ankyrin repeat homology <AN22>
F;766-798/Domain: ankyrin repeat homology <AN23>

alignment_scores:

Quality: 290.00 Length: 230
Ratio: 1.959 Gaps: 3
Percent Similarity: 64.348 Percent Identity: 34.348

alignment_block:

US-09-509-775-1 x B35049

Align seg 1/1 to: B35049 from: 1 to: 1856

109 GTGTCTAACCTTAATGGTCTGCAACCTGGCCTACAGCGGAA 149
||||| |||
432 ValSerAsnValLysValGluThrProLeuHisMetAlaAlaArgAlaG 448
150 GCTGGAGAGTTGAAGGAGAGATATTCTGGCGGATAAATCCCTGGCTACTA 199
||||| |||
448 yHisThrGluValAlaLysTyrLeuLeuGlnAsnLysAlaLysValAsnA 465
200 GAACTGACCGAGCAGCAGCACTGCTTGCATGCGCATGCTCAGCTGGA 249
||||| |||
465 laLysAlaLysAspAspGlnThrProLeuHisCysAlaAlaArgIleGly 481
250 CATACAGAAATTTGTAATTTTGTTCGAATTTGTCAGTGGCAGTGAATGA 299
||||| |||
482 HisThrAsnMetValLysLeuLeuGluAsnAlaAsnProAsnLe 498
300 TAAAGACGATGCGAGTTGGTCTCTCTCTCATATTGCGGCTTCTCTGCCC 349
||||| |||
498 uAlaThrAlaGlyHisThrProLeuHisIleAlaAlaArgGluGly 515
350 GGGATGAGATTTAAAGCCCTTCTGGGAAAGGCTCAAGTGAATGCT 399
||||| |||
515 isValGluThrValLeuAlaLeuLeuGluLysGluAlaSerGlnAlaCys 531
400 GTCNATCAAAATGCTGCTACTCTCTCTTACATTTACATTCAGCTCGAAAAACAG 449
||||| |||
532 MetThrLysLysGlyPheThrProLeuHisValAlaAlaLysTyrGlyLy 548
450 GCATGAGATGCTGCTCATGTACTGGAAGGGGCTAATCCAGATGCTA 499
||||| |||
548 sValArgValAlaGluLeuLeuGluArgAspAlaHisProAsnAlaA 565
500 AGGACCATTTATGAGCTACAGCAATGACCGGGCAGCAGCAAGGGTAAC 549
||||| |||
565 laGlyLysAsnGlyLeuThrProLeuHisValAlaValHisHisAsnAsn 581

R;Lux, S.E.; John, K.M.; Bennett, V.
Nature 344, 36-42, 1990
A:Title: Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure
A:Reference number: S08275; MUID:90158830
A:Accession: S08275
A:Molecule type: mRNA
A:Residues: 1-1881 <LUI>
A:Cross-references: EMBL:X16609; NID:g28701; PIDN:CAA34610.1; PID:g28702
A:Accession: A33219
A:Molecule type: protein
A:Residues: 2-7,'X',9-17,'X',19-20,'T',22-30:733-749,'A',751-753:828-833,'X',835-855,'X'
X'1367:1383-1427:1601-1630:1686-1698,'D',1700:1763-1772 <LUX>
A:Note: 845-Arg and 1392-Thr were also found
A:Hermann, J.; Barel, M.; Frade, R.
Biochem. Biophys. Res. Commun. 204, 453-460, 1994
A:Title: Human erythrocyte ankyrin, a cytoskeleton component, generates the p57 membrane
A:Reference number: PC2220; MUID:95071348
A:Accession: PC2220
A:Molecule type: protein
A:Residues: 910-929 <HER>
R;Davis, L.H.; Bennett, V.
J. Biol. Chem. 265, 10589-10596, 1990
A:Title: Mapping the binding sites of human erythrocyte ankyrin for the anion exchanger
A:Reference number: A35443; MUID:90285190
A:Accession: A35443
A:Molecule type: protein
A:Residues: 'X',5,'X',7-12:403-417,'X',419-422,'H',424,'LQ':797-800,'L',802-814:862-863,
C:Genetics:
A:Gene: GDB:ANK1; ANK
A:Cross-references: GDB:118737; OMIM:182900
A:Map position: 8p11.2-8p11.2
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing; phosphoprotein
F:2-1881/Product: ankyrin 1, erythrocyte form 1
F:2-1512,1675-1881/Product: ankyrin 2, erythrocyte form 2
F:2-827/Domain: 89K #status predicted <DOM1>
F:2-827/Region: anion exchange protein binding
F:44-76/Domain: ankyrin repeat homology <AN01>
F:77-109/Domain: ankyrin repeat homology <AN01>
F:110-142/Domain: ankyrin repeat homology <AN02>
F:143-171/Domain: ankyrin repeat homology <AN03>
F:172-204/Domain: ankyrin repeat homology <AN04>
F:205-237/Domain: ankyrin repeat homology <AN05>
F:238-270/Domain: ankyrin repeat homology <AN06>
F:271-303/Domain: ankyrin repeat homology <AN07>
F:304-336/Domain: ankyrin repeat homology <AN08>
F:337-369/Domain: ankyrin repeat homology <AN09>
F:370-402/Domain: ankyrin repeat homology <AN10>
F:403-435/Domain: ankyrin repeat homology <AN11>
F:436-468/Domain: ankyrin repeat homology <AN12>
F:469-501/Domain: ankyrin repeat homology <AN13>
F:502-534/Domain: ankyrin repeat homology <AN14>
F:535-567/Domain: ankyrin repeat homology <AN15>
F:568-600/Domain: ankyrin repeat homology <AN16>
F:601-633/Domain: ankyrin repeat homology <AN17>
F:634-666/Domain: ankyrin repeat homology <AN18>
F:667-699/Domain: ankyrin repeat homology <AN19>
F:700-732/Domain: ankyrin repeat homology <AN20>
F:733-765/Domain: ankyrin repeat homology <AN21>
F:766-798/Domain: ankyrin repeat homology <AN22>
F:828-1382/Domain: ankyrin repeat homology <AN23>
F:828-1382/Region: 62K #status predicted <DOM2>
F:1383-1881/Domain: 55K #status predicted <DOM3>

alignment_scores:		
Quality	290.00	Length: 230
Ratio	1.959	Gaps: 3
Percent Similarity	64.348	Percent Identity: 34.348

alignment_block:

US-09-509-775-1 x SJHUK

seq_name: pir2:A57282

seq_documentation_block:

ankyrin-related protein unc-44 -

C;Species: *Caenorhabditis elegans*

C;Date: 01-Dec-1995 #sequence_rev

C;Accession: A57282

R;Otsuka, A.J.; Franco, R.; Yang, J. *Cell Biol* 129 1081-1083 1990

A;Title: An ankyrin-related gene

A; Reference number: A57282; MUID:

A;Accession: A57282

A;Status: preliminary

A;Molecule type: DNA
A.Residues: 1-1795

A;Residues: 1-1786 <OTS>

A:Cross-references: GB:U21734; NID:G790607; PIDN:AAA85854.1; PID:G790608
C:Genetics:
A:Gene: unc-44
A:Introns: 97/3; 262/2; 535/1; 749/1; 786/3; 883/3; 1001/2; 1007/2; 1051/3; 1666/1; 1725/1
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
F:164-182/Domain: ankyrin repeat homology <AN04>
F:358-390/Domain: ankyrin repeat homology <AN1>
F:391-423/Domain: ankyrin repeat homology <AN11>

alignment_scores:
Quality: 288.50 Length: 211
Ratio: 2.003 Gaps: 4
Percent Similarity: 68.246 Percent Identity: 36.019
alignment_block:
US-09-509-775-1 x A57282 ..
Align seg 1/1 to: A57282 from: 1 to: 1786

```
154 GAAGAGCTTGAAGGAGAGATATTCTG.....GCCGATAAATCCTGCGC 194
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
537 GluGluValAlaGlyIleLeuLeuAspHisAsnAlaAspLysThrLeuLe 553
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
195 TACTAGAACTGACGAGCAGCAGCAACTGCATTGCGACTGGCGTCTCAG 244
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
553 uThrLys.....LysGlyPheThrProLeuHisLeuAlaSerLysT 567
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
245 CTGGACATACAGAAATGTTGAATTTTGTTCGAACCTGGAGTGGCAGTG 294
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
567 yrGlyAsnLeuGluValAlaArgLeuLeuGluGluArgGlyThrProVal 583
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
295 AATGATAAAGACGATCCAGGTTGGTCTCTCTCATATTGCGGCTCTCGC 344
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
584 AspIleGluGlyLysAsnGlnValThrProLeuHisValAlaAlaHisT 600
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
345 TGCCCGGGATGAGATTGTAAGAGCCCTTCTGGGAAAAGGTGCTCAAGTGA 394
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
600 rAsnAsnAspLysValAlaMetLeuLeuLeuGluAsnGlyAlaSerAlaL 617
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
395 ATGCTGTCAATCAAAATGCTGTACCTCTTACATTTATGCGACTTCGAAA 444
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
617 ysAlaAlaAlaLysAsnGlyTyrThrProLeuHisIleAlaAlaLysLys 633
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
445 AACAGCATGAGATCGCTCATGTCTGTTACTGGAAGCGGGCTTAATCCAGA 494
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
634 AsnGlnMetGluIleAlaSerThrLeuLeuGlnPheLysAlaAspProAs 650
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
495 TGCTAAGGACCATTTATGAGGCTACAGCAATGCGCGGCGAGCGCAAGG 544
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
650 nAlaLysSerArgAlaGlyPheThrProLeuHisLeuSerAlaGlnGluG 667
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
545 GTAACCTGAAGATGATTATATCTTCTTCTTACTACAAAGCATCCACAAC 594
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
667 lYHisLysGluIleSerGlyLeuLeuIleGluAsnGlySerAspValGly 683
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
595 ATCCAGACACTGAGGTAAACACTCCTCTACACTTAGCTGTGATCCAGGA 644
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
684 AlaLysAlaAsnAsnGlyLeuThrAlaMetHisLeuCysAlaGlnGluAs 700
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
645 GAGAGTGGGAAGCAAGCAAACTGCTGTGTGTCCTCCCAAGGAGCAAGTAT 694
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
700 pHisValProValAlaGlnIleLeuTyrAsnAsnGlyAlaGluIleAsnS 717
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
695 TTGAGATAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 738
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
717 erLysThrAsnAlaGlyTyrThrProLeuHisValAlaCysHisPheGly 733
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
739 GCGCTGGGTTAATCTCAAGAGATGGTGGAA 771
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
734 GlnLeuAsnMetVal...LysPheLeuValGlu 743
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

seq_name: p1r2:T15346

seq_documentation_block:
elegans ankyrin-related unc-44 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jan-2000
C:Accession: T15346
R:Gattung, S.
submitted to the EMBL Data Library, February 1996
A:Description: The sequence of C. elegans cosmid B0350.
A:Reference number: Z18332
A:Accession: T15346
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1815 <GAT>
A:Cross-references: EMBL:U50071; NID:gl208871; PID:gl208875; PIDN:AAA93445.1; CESP:un
C:Genetics:
A:Gene: CESP:unc-44
A:Introns: 97/3; 262/2; 535/1; 749/1; 786/3; 883/3; 1001/2; 1007/2; 1051/3; 1666/1; 1
C:Superfamily: ankyrin; ankyrin repeat homology
F:358-390/Domain: ankyrin repeat homology <ANR>

alignment_scores:
Quality: 288.50 Length: 211
Ratio: 2.003 Gaps: 4
Percent Similarity: 68.246 Percent Identity: 36.019
alignment_block:
US-09-509-775-1 x T15346 ..
Align seg 1/1 to: T15346 from: 1 to: 1815

```
154 GAAGAGTTGAAGGAGAGATATTCTG.....GCCGATAAATCCTGCGC 194
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
537 GluGluValAlaGlyIleLeuLeuAspHisAsnAlaAspLysThrLeuLe 553
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
195 TACTAGAACTGACGAGCAGCAGCAACTGCATTGCGACTGGCGTCTCAG 244
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
553 uThrLys.....LysGlyPheThrProLeuHisLeuAlaSerLysT 567
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
245 CTGGACATACAGAAATGTTGAATTTTGTTCGAACCTGGAGTGGCAGTG 294
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
567 yrGlyAsnLeuGluValAlaArgLeuLeuGluGluArgGlyThrProVal 583
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
295 AATGATAAAGACGATCCAGGTTGGTCTCTCTCATATTGCGGCTCTCGC 344
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
584 AspIleGluGlyLysAsnGlnValThrProLeuHisValAlaAlaHisT 600
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
345 TGCCCGGGATGAGATTGTAAGAGCCCTTCTGGGAAAAGGTGCTCAAGTGA 394
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
600 rAsnAsnAspLysValAlaMetLeuLeuLeuGluAsnGlyAlaSerAlaL 617
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
395 ATGCTGTCAATCAAAATGCTGTACCTCTTACATTTATGCGACTTCGAAA 444
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
617 ysAlaAlaAlaLysAsnGlyTyrThrProLeuHisIleAlaAlaLysLys 633
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
445 AACAGCATGAGATCGCTCATGTCTGTTACTGGAAGCGGGCTTAATCCAGA 494
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
634 AsnGlnMetGluIleAlaSerThrLeuLeuGlnPheLysAlaAspProAs 650
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
495 TGCTAAGGACCATTTATGAGGCTACAGCAATGCGCGGCGAGCGCAAGG 544
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
650 nAlaLysSerArgAlaGlyPheThrProLeuHisLeuSerAlaGlnGluG 667
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
545 GTAACCTGAAGATGATTATATCTTCTTCTTACTACAAAGCATCCACAAC 594
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
667 lYHisLysGluIleSerGlyLeuLeuIleGluAsnGlySerAspValGly 683
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
595 ATCCAGACACTGAGGTAAACACTCCTCTACACTTAGCTGTGATCCAGGA 644
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
684 AlaLysAlaAsnAsnGlyLeuThrAlaMetHisLeuCysAlaGlnGluAs 700
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
645 GAGAGTGGGAAGCAAGCAAACTGCTGTGTGTCCTCCCAAGGAGCAAGTAT 694
```

700 phisValProValAlaGlnIleLeuTyrAsnAsnGlyAlaGluIleAsnS 717
:::||||| |||||:::||||| ::::::|||:::|||||
695 TTGAATAAAGAGAAGAAACCCCTCGAATGCC.....AAAGGT 738
:::||||| |||||:::||||| |||||
717 eLysThrAsnAlaGlyTyrThrProLeuHisValAlaCysHisPheGly 733
:::||||| |||||:::||||| |||||
739 GCCTGGGTTTAATACTCAAGAGAANTGGTGGA 771
|||||:::||||| |||||:::|||||
734 GlnLeuAsnMetVal...LysPheLeuValGlu 743

THIS PAGE BLANK (USPTO)

OM of: US-09-509-775-1 to: SwissProt_39.* out_format : pfs
Date: Aug 13, 2001 8:37 AM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+np.model -DRV=xlp
-Q/cgn2_1/USPTO_spool/US09509775/runat_13082001_074528_13094/app_query.fasta_1.848
-DB=SwissProt_39 -QFMT=fastan -SUFFIX=rs -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blasum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM_ext -MINLEN=0 -MAXLEN=200000000
-USER=US09509775 -CGN1_1_28 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:

Query: US-09-509-775-1
Query length: 780
Database: SwissProt_39.*
Database sequences: 33435
Database length: 3425486
Search time (sec): 62.940000

score_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
SwissProt_39:ANK2_HUMAN	+	294.00	456.18	2.3e-19	3924
SwissProt_39:Y64X_YEAST	+	291.00	475.54	3.3e-19	228
SwissProt_39:ANK1_HUMAN	+	290.00	455.83	5.1e-19	1880
SwissProt_39:ANK1_MOUSE	+	287.50	451.76	8.1e-19	1862
SwissProt_39:YD57_SCHPO	+	273.50	446.24	1.4e-17	234
SwissProt_39:PA26_MOUSE	+	254.50	404.68	9.0e-16	732
SwissProt_39:ANR3_HUMAN	+	248.50	393.85	3.2e-15	832
SwissProt_39:PA26_RAT	+	247.50	393.06	4.0e-15	751
SwissProt_39:LATA_LATMA	+	247.50	387.73	4.2e-15	1401
SwissProt_39:DAPK_HUMAN	+	247.00	386.71	4.7e-15	1431
SwissProt_39:PA26_HUMAN	+	237.50	375.84	3.4e-14	806
SwissProt_39:AKR_ATHAT	+	236.50	379.38	3.9e-14	439
SwissProt_39:ANKH_CHRVI	+	225.00	362.90	4.4e-13	333
SwissProt_39:KBF2_HUMAN	+	224.50	353.32	5.4e-13	898
SwissProt_39:ILK_MOUSE	+	217.50	347.56	2.3e-12	452
SwissProt_39:PEM1_CAEEL	+	215.00	340.22	4.0e-12	656
SwissProt_39:NOTC_DROME	+	215.00	328.10	4.6e-12	2703
SwissProt_39:NTC3_MOUSE	+	214.00	327.76	5.6e-12	2318
SwissProt_39:ILK_CAVPO	+	212.50	339.27	6.6e-12	451
SwissProt_39:NTC1_RAT	+	212.00	323.68	8.7e-12	2531
SwissProt_39:ILK1_HUMAN	+	211.50	337.59	8.1e-12	452
SwissProt_39:ILK2_HUMAN	+	211.50	337.59	8.1e-12	452
SwissProt_39:NTC1_HUMAN	+	209.50	319.83	1.5e-11	2444
SwissProt_39:NTC4_MOUSE	+	208.50	320.03	1.8e-11	1964
SwissProt_39:NOTC_BRARE	+	207.50	316.53	2.2e-11	2437
SwissProt_39:RN5A_MOUSE	+	202.00	338.32	6.4e-11	679
SwissProt_39:BCL3_HUMAN	+	201.00	320.26	7.6e-11	446
SwissProt_39:NOTC_XENLA	+	200.00	303.76	1.1e-10	2524
SwissProt_39:KBF1_CHICK	+	198.00	308.50	1.6e-10	984
SwissProt_39:RN5A_HUMAN	+	194.00	304.28	3.6e-10	741
SwissProt_39:AKR1_YEAST	+	194.00	304.02	3.6e-10	764
SwissProt_39:GABB_HUMAN	+	189.50	303.30	8.6e-10	347
SwissProt_39:GABB_MOUSE	+	189.50	302.45	8.7e-10	383
SwissProt_39:KBF1_HUMAN	+	188.00	292.02	1.3e-09	959
SwissProt_39:KBF1_MOUSE	+	187.50	296.85	1.4e-09	500
SwissProt_39:KBF1_MOUSE	+	187.50	291.17	1.5e-09	971
SwissProt_39:L112_CAEEL	+	187.00	287.03	1.7e-09	1429
SwissProt_39:GABB_MOUSE	+	186.50	297.49	1.6e-09	382
SwissProt_39:KBF1_RAT	+	186.00	293.99	1.9e-09	522

SwissProt_39:NTC1_MOUSE + 183.50 276.33 3.8e-09 2531
SwissProt_39:CDN6_HUMAN + 183.00 298.71 3.2e-09 168
SwissProt_39:HT16_HYDAT + 182.50 284.99 4.1e-09 757
SwissProt_39:YA2A_SCHPO + 181.50 284.74 5.0e-09 642
SwissProt_39:RFXK_HUMAN + 181.00 291.65 5.1e-09 260

seq_name: SwissProt_39:ANK2_HUMAN

seq_documentation_block:

AD ANK2_HUMAN STANDARD; PRT; 3924 AA.
AC Q01484; Q01485;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ANKYRIN 2 (BRAIN ANKYRIN) (ANKYRIN B) (ANKYRIN, NONERYTHROID).
GN ANK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain stem;
RX MEDLINE=94075409; PubMed=8253844;
RA Chan W., Kordeli E., Bennett V.;
RT "440-KD ankyrinB: structure of the major developmentally regulated
domain and selective localization in unmyelinated axons.";
RL J. Cell Biol. 123:1463-1473(1993).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RC TISSUE=Brain stem;
RX MEDLINE=91302466; PubMed=18330053;
RA Otto E., Kunimoto M., McLaughlin T., Bennett V.;
RT "Isolation and characterization of cDNAs encoding human brain
ankyrins reveal a family of alternatively spliced genes.";
RL J. Cell Biol. 114:241-253(1991).
RN [3]
RP REVISIONS.
RA Carpenter S.;
RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 463-495 FROM N.A.
RX MEDLINE=92009921; PubMed=1833308;
RA Tse W.T., Menninger J.C., Yang-Feng T.L., Francke U., Sahr K.E.,
Lux S.E., Ward D.C., Forget B.G.;
RT "Isolation and chromosomal localization of a novel nonerythroid
ankyrin gene.";
RL Genomics 10:858-866(1991).
CC -1- FUNCTION: ANKYRINS ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTO-
SKELETAL ELEMENTS; THEY BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN
BAND 4.2, TO NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85,
AND TO THE CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND
DESMIN. ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO
THE CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;
THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3; ARE
PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: PLASMA MEMBRANE OF NEURONS AS WELL AS GLIAL
CELLS THROUGHOUT THE BRAIN.
CC -1- PTM: PHOSPHORYLATED AT MULTIPLE SITES BY DIFFERENT PROTEIN KINASES
AND EACH PHOSPHORYLATION EVENT REGULATES THE PROTEIN'S STRUCTURE
AND FUNCTION (POTENTIAL).
CC -1- SIMILARITY: CONTAINS 23 ANK REPEATS.
CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
or send an email to license@isb-sib.ch).
CC -----

```
DR EMBL: Z26634; CAB42644.1; -
DR EMBL: X56957; CAA40278.1; -
DR EMBL: X56958; CAA40279.2; -
DR EMBL: M37123; AAG62828.1; -
DR PIR: S14533; S14533.
DR PIR: A39643; A39643.
DR PIR: B39643; B39643.
DR PIR: S14569; S14569.
DR HSSP: Q00420; LANC.
DR MIM: 106410; -.
DR InterPro: IPR000488; -.
DR InterPro: IPR000906; -.
DR InterPro: IPR002110; -.
DR Pfam: PF00791; Z05; 1.
DR Pfam: PF00023; ank; 22.
DR Pfam: PF00531; death; 1.
DR PROSITE: PS50088; ANK_REPEAT; 20.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS50017; DEATH_DOMAIN; 1.
DR Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
KW Phosphorylation; Multigene family.
FT REPEAT 63 92
FT REPEAT 96 125 ANK 2.
FT REPEAT 129 158 ANK 3.
FT REPEAT 162 191 ANK 4.
FT REPEAT 193 220 ANK 5.
FT REPEAT 232 261 ANK 6.
FT REPEAT 265 294 ANK 7.
FT REPEAT 298 327 ANK 8.
FT REPEAT 331 360 ANK 9.
FT REPEAT 364 393 ANK 10.
FT REPEAT 397 426 ANK 11.
FT REPEAT 430 459 ANK 12.
FT REPEAT 463 492 ANK 13.
FT REPEAT 496 525 ANK 14.
FT REPEAT 529 558 ANK 15.
FT REPEAT 562 591 ANK 16.
FT REPEAT 595 624 ANK 17.
FT REPEAT 628 657 ANK 18.
FT REPEAT 661 690 ANK 19.
FT REPEAT 694 723 ANK 20.
FT REPEAT 727 756 ANK 21.
FT REPEAT 760 789 ANK 22.
FT REPEAT 793 822 ANK 23.
FT DOMAIN 1773 1950 REPEAT-RICH REGION.
FT REPEAT 1773 1784 REPEAT A.
FT REPEAT 1785 1796 REPEAT A.
FT REPEAT 1797 1808 REPEAT A.
FT REPEAT 1809 1820 REPEAT A.
FT REPEAT 1821 1832 REPEAT A.
FT REPEAT 1833 1844 REPEAT A.
FT REPEAT 1845 1856 REPEAT A.
FT REPEAT 1857 1867 REPEAT A. (APPROXIMATE).
FT REPEAT 1868 1879 REPEAT A.
FT REPEAT 1880 1891 REPEAT A.
FT REPEAT 1892 1902 REPEAT A. (APPROXIMATE).
FT REPEAT 1903 1914 REPEAT A.
FT REPEAT 1915 1926 REPEAT A.
FT REPEAT 1927 1938 REPEAT A.
FT REPEAT 1939 1950 REPEAT A.
FT DOMAIN 3536 3620
FT VARSPPLIC 1039 1039
FT VARSPPLIC 1444 3528
FT VARSPPLIC 475 476 MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT CONFLICT 971 971 GO -> PE (IN REF. 4).
FT CONFLICT 3581 3582 I -> S (IN REF. 2).
FT CONFLICT 3586 3586 QY -> HA (IN REF. 2).
FT CONFLICT 3586 3586 I -> Y (IN REF. 2).
FT SEQUENCE 3924 AA; 430337 MW; 52AC496C428E29D2 CRC64;
```

alignment_scores:

Quality: 294.00

Length: 232

```
Ratio: 2.146 Gaps: 2
Percent Similarity: 59.052 Percent Identity: 31.897
alignment_block:
US-09-509-775-1 x ANK2_HUMAN
Align seg 1/1 to: ANK2_HUMAN from: 1 to: 3924
136 GCCTACAGCGGGGAAGCTGGAAGAGTTCAGAGGAGTAGTATTGCGCCGATAA 185
||| :|||:||||| :|||: :|||: :|||: :|||: :|||: :|||: :|||:
472 AlaArgAlaGlyGlnValValArg...CysLeuLeuArgAsnG1 487
186 ATCCCTGGCTACTAGAACTGACAGGAGCAGCAGAACTGCATTCACACTGGG 235
:|||||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
487 yAlaLeuValAspAlaArgAlaArgGluGluGlnThrProLeuHisIleA 504
236 CATGCTCAGCTGCACATACAGAAATTTGCAATTTTGTTCGAA..... 279
||| :|||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
504 laSerArgLeuGlyLysThrGluIleValGlnLeuLeuGlnHisMet 520
279 ..... 279
521 AlaHisProAspAlaAlaThrThrAsnGlyTyrThrProLeuHisIleSe 537
280 .....CTTGGAG 286
537 rAlaArgGluGlyGlnValAspValAlaSerValLeuLeuGluAlaGlyA 554
287 TGCCAGTGAATGATAAGAGCAGTGCAGGTGGTCTCTCTTCATATTCG 336
: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
554 laAlaHisSerLeuAlaThrLysLysGlyPheThrProLeuHisValAla 570
337 GCTTCTGCTGGCGGATGAGATTGTAAGAGCCCTTCTGGGAAAGGTGC 386
|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
571 AlaLysTyrGlySerLeuAspValAlaLysLeuLeuGlnArgArgAl 587
387 TCAAGTGAATGCTCAATCAAAATGGCTGTACTCCCTTACATATTCAG 436
| :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
587 aAlaAlaAspSerAlaGlyLysAsnGlyLeuThrProLeuHisValAla 604
437 CTTGCGAAAACAGCATGAGATCGCTGTCTACTGTGGAAGCGGGGCT 486
|| :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
604 laHisTyrAspAsnGlnLysValAlaLeuLeuLeuGluLysGlyAla 620
487 AATCCAGATGCTAAGGACCATATGAGGCTACAGCAATGACCGGCGAGC 536
:|||||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
621 SerProHisAlaThrAlaLysAsnGlyTyrThrProLeuHisIleAla 637
537 AGCCAGGGTAACTTGAAGATGATTTCATCTCTGTACTACAAAGCAT 586
| :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
637 aLysLysAsnGlnMetGlnIleAlaSerThrLeuLeuAsnTyrGlyAla 654
587 CCACAACATCCAGACACGTAGGGAAGCACTCTCTACACTTAGCCTGT 636
:|||||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
654 luThrAsnIleValThrLysGlnGlyValThrProLeuHisLeuAlaSer 670
637 GATCAGGAGAGTGGGAAGCAAACTGCTGTGTCCCAAGGAGCAAG 686
:|||||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
671 GlnGluGlyHisThrAspMetValThrLeuLeuAspLysGlyAlaAs 687
687 TATTTCATTGAGAAATAAGAAAGAACACCCCTGCAAGTGGCC 732
:|||||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
687 nIleHisMetSerThrLysSerGlyLeuThrSerLeuHisLeuAla 702
seq_name: SwissProt_39:YG4X_YEAST
seq_documentation_block:
ID YG4X_YEAST STANDARD; PRT; 228 AA.
AC P50086;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HYPOTHETICAL 25.6 KDA PROTEIN IN SM11-PHO81 INTERGENIC REGION.
```


[illegible]

FT ACT_SITE 139 139 BY SIMILARITY.
FT MUTAGEN 42 42 K->A: LOSS OF ACTIVITY.
SQ SEQUENCE 1431 AA; 160017 MW; 9EB84811004A155B CRC64;

alignment_scores:

Quality: 247.00 Length: 210
Ratio: 2.111 Gaps: 1
Percent Similarity: 55.714 Percent Identity: 29.048

alignment_block:

US-09-509-775-1 x DAPK_HUMAN

Align seg 1/1 to: DAPK_HUMAN from: 1 to: 1431

205 GACCAGGACAGCAGACTCCATTGACCTGGGATGCTCAGCTGGACATAC 254
||||| : : : : : ||||| : : : : : ||||| : : : : :
409 ASPLYSGLYGLYSERASNAVALTYRTPALAAARGHISGLYHISVA 425
: : : : : ||||| : : : : : ||||| : : : : :
255 AGAAATGTTGAATTTTGTGCAACTTGGAGTGGCAGTGAATGATAAAG 304
: : : : : ||||| : : : : : ||||| : : : : :
425 IASPTHRLEULYSPEHLEUSERGLUASNLYSCYSPROLEUASPVALLYSA 442
: : : : : ||||| : : : : : ||||| : : : : :
305 AGCATGACAGTGGTCTCTTCATATGTCGGCTTCTGCTGCGCGGAT 354
|| : : ||| ||||| : : ||||| : : ||||| : : ||||| : :
442 SPLYSERGLYGLUMETALALEUHHISVALAALAARGTYRGLYHISALA 458
: : : : : ||||| : : : : : ||||| : : : : :
355 GAG..... 357
: : : : :
459 ASPVALAALGVALTHRCSYSAALASERALAGINLEPROILESERAR 475
357 357
475 GTHRYSGLUGLUGLUTHRPROLEUHHISCSYSAALATRPHISGLYTYR 492
358ATTGTAAGCCCTTCTGGGAAAGGTGCTCAAGTGAATGCTGTC 402
: : : : : ||||| : : : : : ||||| : : : : :
492 YR-SERVALALALYSALALEUCYSGLUALAGLYCYSASNVALASNILEY 508
: : : : : ||||| : : : : : ||||| : : : : :
403 AATCAAAATGGCTGACTCCCTTACATATGTCAGCTTCGAAAAACAGGCA 452
||||| : : : : : ||||| : : : : : ||||| : : : : :
509 ASNARGGLUGLUGLUTHRPROLEUHLTHRALASERALARGGLYTYRHI 525
: : : : : ||||| : : : : : ||||| : : : : :
453 TGAGATCGCTGTCATGTTACTGGAAGCGGGCTAATCCAGATGCTAAGG 502
: : : : : ||||| : : : : : ||||| : : : : :
525 SASPILEVALGLUCYSLEUALAGLHISGLYALASPHEUASNALACYSAA 542
: : : : : ||||| : : : : : ||||| : : : : :
503 ACCATTATGAGGCTACAGCAATGCCACCGGCGAGCCAGCGGTAACTTG 552
|| : : : : : ||||| : : : : : ||||| : : : : :
542 SPLYSASPOLYHISTILEALALEUHHISLEUALAVALARGCYSGLINMET 558
: : : : : ||||| : : : : : ||||| : : : : :
553 AAGATGATTCATATCTCTTCTGCTACTACAAAGCATCCACAAACATCAAGA 602
: : : : : ||||| : : : : : ||||| : : : : :
559 GLUVALILELYSLEUTHRLEUHLTHRALASERALARGGLYTYRGLNAS 575
603 CACTGAGGTAACTCTCTTACACTTAGCTGATGAGGAGAGAGTGG 652
| : : : : ||||| : : : : : ||||| : : : : :
575 PARGLHISGLYASNTHRPROLEUHHISVALAALACYSGLYASNMECP 592
: : : : : ||||| : : : : : ||||| : : : : :
653 AAGAAGAAACAACTGCTGTGTGCTCCCAAGGAGCAAGTATTACATGAGAT 702
: : : : : ||||| : : : : : ||||| : : : : :
592 TOILEVALVALALALEUCYSGLUALASNCYSASNLEUASPILASERASN 608
703 AAGAAGAAACAAACCTCTGCAAGTGGCC 732
||| : : : ||||| : : : : : ||||| : : : : :
609 LYSTYRGLYARGTHRPROLEUHHISLEUALA 618

seq_name: SwissProt_39:PA26_HUMAN

seq_documentation_block:

ID PA26_HUMAN STANDARD; PRT; 806 AA.
AC O60733; O75645; O9Y671; O9UIT0; Q9UG29;
DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 01-OCN-2000 (Rel. 40, Last annotation update)
DE 85 KDA CALCIUM-INDEPENDENT PHOSPHOLIPASE A2 (EC 3.1.1.4) (IPLA2) (CAI-
DE PLA2).
GN PLA2G6 OR IPLA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS LH-IPLA; ANKYRIN-IPLA2-1 AND -2).
RC TISSUE-B-cell, and Testis;
RX MEDLINE=98079046; PubMed=9417066;
RA Larsson P.K.A., Claesson H.-E., Kennedy B.P.;
RT "Multiple splice variants of the human calcium-independent
RT phospholipase A2 and their effect on enzyme activity.";
RL J. Biol. Chem. 273:207-214(1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS LH-IPLA AND SH-IPLA).
RC TISSUE-Pancreatic islets;
RX MEDLINE=99194813; PubMed=10092647;
RA Ma Z., Wang X., Nowatzke W., Ramanadham S., Turk J.;
RT "Human pancreatic islets express mRNA species encoding two distinct
RT catalytically active isoforms of group VI phospholipase A2 (IPLA2)
RT that arise from an exon-skipping mechanism of alternative splicing of
RT the transcript from the IPLA2 gene on chromosome 22q13.1.";
RL J. Biol. Chem. 274:9607-9616(1999).
RN [3]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=99269033; PubMed=10336645;
RA Larsson Forsell P.K.A., Kennedy B.P., Claesson H.-E.;
RT "The human calcium-independent phospholipase A2 gene. Multiple enzymes
RT with distinct properties from a single gene.";
RL Eur. J. Biochem. 262:575-585(1999).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM LH-IPLA).
RC Graffham D.;
RX Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM LH-IPLA).
RC TISSUE-Testis;
RX Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RA Ansorge W., Wirkner U., Mewes H.-W., Gassenhuber J., Wiemann S.;
RT IT HAS BEEN IMPLICATED IN NORMAL PHOSPHOLIPID REMODELLING, NITRIC
RT OXIDE-INDUCED OR VASOPRESSIN-INDUCED ARACHIDONIC ACID RELEASE AND
RT IN LEUKOTRIENE AND PROSTAGLANDIN PRODUCTION. MAY PARTICIPATE IN
RT FAS MEDIATED APOPTOSIS AND IN REGULATING TRANSMEMBRANE ION FLUX IN
RT GLUCOSE-STIMULATED B-CELLS.
CC -1- FUNCTION: ISOFORMS ANKYRIN-IPLA2-1 AND -2, WHICH LACK THE
CC CATALYTIC DOMAIN, ARE PROBABLY INVOLVED IN THE NEGATIVE REGULATION
CC OF IPLA2 ACTIVITY.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERYL-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- SUBUNIT: FORMS LARGE OLIGOMERIC 270-350 KDA STRUCTURES.
CC -1- SUBCELLULAR LOCATION: THE LONG ISOFORM WAS FOUND TO BE MEMBRANE
CC BOUND. THE SHORT ISOFORM IS CYTOPLASMIC.
CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; LH-IPLA2 (SHOWN HERE), SH-IPLA2,
CC ANKYRIN-IPLA2-1 AND ANKYRIN-IPLA2-2; ARE PRODUCED BY ALTERNATIVE
CC SPLICING.
CC -1- TISSUE SPECIFICITY: FOUR DIFFERENT TRANSCRIPTS WERE FOUND TO BE
CC EXPRESSED IN A DISTINCT TISSUE DISTRIBUTION.
CC -1- SIMILARITY: CONTAINS 7 ANK REPEATS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF064594; AAC97486.1; -.

RA MEDLINE-93104681; PubMed-1281700;
RT Zhang H., Scheiner D.C., Fowle W.H., Goodman H.M.;
RT "Expression of antisense or sense RNA of an ankyrin repeat-containing
RT gene blocks chloroplast differentiation in Arabidopsis";
RL Plant Cell 4:1575-1588(1992).
CC -1- FUNCTION: MAY PLAY A ROLE IN THE TEMPORAL AND SPATIAL REGULATION
CC OF CHLOROPLAST DEVELOPMENT FROM PROPLASTID.
CC -1- DEVELOPMENTAL STAGE: HIGHEST EXPRESSION OCCURS IN TWO-WEEK-OLD
CC PLANTS AND DECLINES AS PLANTS DEVELOP FURTHER.
CC -1- SIMILARITY: CONTAINS 4 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M82883; AAA32812.1; -
DR PIR; J01729; J01729.
DR HSSP; Q13625; 1YCS.
DR InterPro; IPR002110; -
DR Pfam; PF00023; ank; 4.
DR PROSITE; PS50088; ANK_REPEAT; 3.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW Cytoskeleton; Repeat; ANK repeat; Multigene family.
FT REPEAT 288 317 ANK 1.
FT REPEAT 321 350 ANK 2.
FT REPEAT 354 383 ANK 3.
FT REPEAT 387 416 ANK 4.
SQ SEQUENCE 439 AA; 49150 MW; C371A90028B25BF3 CRC64;

alignment_scores:

Quality: 236.50 Length: 161
Ratio: 2.319 Gaps: 2
Percent Similarity: 63.354 Percent Identity: 37.267

alignment_block:

US-09-509-775-1 x AKR_ARATH ..
Align seg 1/1 to: AKR_ARATH from: 1 to: 439
265 GAATTTTCTTG.....CAACTTGGAGTCCCAATGA 299
272 LysPheMetLeuAsnSerArgAsnProAspLeuAlaValAlaThrSerLy 288
300 TAAAGACGATGCAGGTGGCTCTCTCATATTGGCGTCTCGCTGGCC 349
288 slys.....TrpLeuProLeuHisThrLeuAlaAlaCysGlyC 301
350 GGGATGAGATGTTAAAGCCCTTCTGGGAAAGGTGCTCAAGTGAATGCT 399
301 lrpPheTyrLeuValAspSerLeuLeuLysHisAsnLeuAspLeuAla 317
400 GTCATCAAAATGGCTGTACTCCCTTACATTATGCAGCTTCGNAACACAG 449
318 ThrAspValGlyGlyLeuThrAlaLeuHisArgAlaAlaIleGlyLysLy 334
450 GCATGAGATCGCTGTCTATGTTACTGGAAGCGGGGCTAATCCAGATGCTA 499
334 sGlnAlaIleThrAsnTyrLeuLeuArgGluSerAlaAsnProPheVal 351
500 AGGACATTATGAGTCAGCAGTACCGCGGCGCAGCAGCCAGCGGTAAC 549
351 euAspAspGluGlyAlaThrLeuMetHisTyrAlaValGlnThrAlaSer 367
550 TTGAAGATGATTCATCTCTCTGTACTACAAGCATCCACAACATCCA 599
368 AlaProThrIleLysLeuLeuLeuTyrAsnAlaAspIleAsnAlaG 384
600 AGACACTGAGGGTAAACACTCCTCTACACTTAGCCTGTGATGAGGAGAG 649

384 naspArgAspGlyTyrThrProLeuHisValAlaValGlnAlaArgS 401
650 TCGAAGAACAACAACTGGTCTCCCAAGGCAAGTATTACATTTGAG 699
401 erAspIleValLysLeuLeuLeuLysGlyAlaAspIleGluValLys 417
700 AATAAAGAAAGAACACCCCTGCAAGTGGCC 732
418 AsnLysAspGlyLeuThrProLeuGlyPheAla 428

seq_name: SwissProt_39:ANKH_CHRVI

seq_documentation_block:

ID ANKH_CHRVI STANDARD; PRT; 323 AA.
AC Q06527;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ANKYRIN HOMOLOG PRECURSOR.
OS Chromatium vinosum.
OC Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae;
OC Allochrochromatium.
OX NCBI_TaxID=1049;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D / ATCC 17899 / DSM 180;
RX MEDLINE-93300842; PubMed-8390993;
RA Dolata M.M., van Beeumen J.J., Ambler R.P., Meyer T.E.,
RA Cusanovich M.A.;
RT "Nucleotide sequence of the heme subunit of flavocytochrome c from
RT the purple phototrophic bacterium, Chromatium vinosum. A 2.6-kilobase
RT pair DNA fragment contains two multiheme cytochromes, a flavoprotein,
RT and a homolog of human ankyrin.";
RL J. Biol. Chem. 268:14426-14431(1993).
CC -1- SIMILARITY: CONTAINS 8 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

CC EMBL; L13419; AAA23315.1; -
CC HSSP; Q00420; IAWC.
CC InterPro; IPR002110; -
CC Pfam; PF00023; ank; 8.
CC PROSITE; PS50088; ANK_REPEAT; 7.
CC PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW Signal; ANK repeat; Repeat.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 323 ANKYRIN HOMOLOG.
FT REPEAT 31 60 ANK 1.
FT REPEAT 64 93 ANK 2.
FT REPEAT 97 126 ANK 3.
FT REPEAT 130 159 ANK 4.
FT REPEAT 163 192 ANK 5.
FT REPEAT 196 225 ANK 6.
FT REPEAT 229 258 ANK 7.
FT REPEAT 262 291 ANK 8.
SQ SEQUENCE 323 AA; 33453 MW; 3976D34A9BD607DF CRC64;

alignment_scores:

Quality: 225.00 Length: 219
Ratio: 1.891 Gaps: 2
Percent Similarity: 54.338 Percent Identity: 29.680

alignment_block:

US-09-509-775-1 x ANKH_CHRVI

seq_documentation_block:				
ID	KB22_HUMAN	STANDARD;	PRT;	898 AA.
AC	Q00653;			
DT	01-OCT-1993	(Rel. 27, Created)		
DT	01-OCT-1996	(Rel. 34, Last sequence update)		
DT	01-OCT-2000	(Rel. 40, Last annotation update)		
DE	LYT10) [CONTAINS: NUCLEAR FACTOR NF-KAPPA-B P100 SUBUNIT (H2TF1) (ONCOGENE LYT-10) (NFkB2).			
GN	NFkB2			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			

[1] SEQUENCE FROM N.A.
RN TSSUB=Leukemia;
RP MEDLINE=91343004; PubMed=1876189;
RX Schmid R.M., Perkins N.D., Duckett C.S., Andrews P.C., Nabel G.J.;
RA "Cloning of an NF-kappa B subunit which stimulates HIV transcription
RT in synergy with p65";
RL Nature 352:733-736(1991).
[2]
RN SEQUENCE FROM N.A.
RP MEDLINE=92123193; PubMed=1531086;
RX Bours V., Burd P.R., Brown K., Villalobos J., Park S., Ryseck R.P.,
RA Bravo R., Kelly K., Siebenlist U.;
RT "A novel mitogen-inducible gene product related to p50/p105-NF-kappa
RL B participates in transactivation through a kappa B site";
RN Mol. Cell. Biol. 12:685-695(1992).
[3]
RN SEQUENCE FROM N.A.
RP MEDLINE=94309910; PubMed=8036016;
RX Thakur S., Lin H.C., Tseng W.T., Kumar S., Bravo R., Foss F.,
RA Gellinas C., Rabson A.B.;
RT "Rearrangement and altered expression of the NFKB-2 gene in human
RL cutaneous T-lymphoma cells";
RN Oncogene 9:2335-2344(1994).
[4]
RN SEQUENCE OF 1-220 FROM N.A.
RP MEDLINE=95059001; PubMed=7969113;
RX Liptay S., Schmid R.M., Nabel E.G., Nabel G.J.;
RA "Transcriptional regulation of NF-kappa B2: evidence for kappa B-
RT mediated 'positive and negative autoregulation.'";
RL Mol. Cell. Biol. 14:7695-7703(1994).
[5]
RN X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 37-327.
RP MEDLINE=98046037; PubMed=9384586;
RX Cramer P., Larson C.J., Verdine G.L., Mueller C.W.;
RA "Structure of the human NF-kappaB p52 homodimer-
RL kappab p52 homodimer-NuA complex at 2.1-A
RN resolution";
RP EMBO J. 16:7078-7090(1997).
[6]
RN FUNCTION: P100 IS THE PRECURSOR OF THE P52 SUBUNIT OF THE NUCLEAR
RP FACTOR NF-KAPPA-B, WHICH BINDS TO THE KAPPA-B CONSENSUS SEQUENCE
RX 5'-GGGNNYYCC-3', LOCATED IN THE ENHANCER REGION OF GENES INVOLVED
RA IN IMMUNE RESPONSE AND ACUTE PHASE REACTIONS. THE PRECURSOR
RL PROTEIN ITSELF DOES NOT BIND TO DNA.
RN SUBUNIT: ACTIVE NF-KAPPA-B IS A HETERODIMER OF AN ABOUT 52 KDA
RP KDA-BINDING SUBUNIT AND THE WEAK DNA-BINDING SUBUNIT P65. TWO
RX HETERODIMERS MIGHT FORM A LABILE TETRAMER.
RA SUBCELLULAR LOCATION: NUCLEAR, BUT ALSO FOUND IN THE CYTOPLASM
RL IN AN INACTIVE FORM COMPLEXED TO AN INHIBITOR (1-KAPPA-B).
RN ALTERNATIVE PRODUCTS: NFKB2 CAN PRODUCE BOTH A 100 KDA AND A 49
RP KDA PRODUCT FROM AN ALTERNATIVELY SPLICED TRANSCRIPT.
RX DOMAIN: THE C-TERMINUS OF P100 MIGHT BE INVOLVED IN CYTOPLASMIC
RA RETENTION, INHIBITION OF DNA-BINDING BY P52 HOMODIMERS, AND/OR
RL TRANSCRIPTION ACTIVATION (BY SIMILARITY).
RN DISEASE: IN CUTANEOUS T-CELL LEUKEMIA (CTCL) CELL LINE HUT 78, A
RP REARRANGED NFKB2 GENE IS FOUND WHICH ENCODES FOR A TRUNCATED 80
RX KDA PROTEIN (P80HT).
RN SIMILARITY: BELONGS TO THE REL/DORSAL FAMILY.
RP SIMILARITY: CONTAINS 7 ANK REPEATS.
[7]
RN THIS SWISS-PROT entry is copyright. It is produced through a collaboration
RP between the Swiss Institute of Bioinformatics and the EMBL outstation -
RX the European Bioinformatics Institute. There are no restrictions on its
RA use by non-profit institutions as long as its content is in no way
RL modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
RP or send an email to license@isb-sib.ch).
[8]
RN EMBL; X61498; CAA33715.1; -;
RP EMBL; S76638; AAB21124.1; -;
RX EMBL; U09609; AAA21462.1; -;
RA EMBL; U20816; AAA68171.1; -;
RL PIR; S17233; S17233.
RN PDB; 1A30; 27-MAY-98.

```

DR TRANSFAC; T01927; -
DR MIM; 164012; -
DR InterPro; IPR000451; -
DR InterPro; IPR000488; -
DR InterPro; IPR002110; -
DR Pfam; PF00554; RHD; 1.
DR Pfam; PF00023; ank; 6.
DR PRINTS; PR00531; death; 1.
DR PROSITE; PRO0057; NFKBNSCPCT.
DR PROSITE; PS50088; ANK_REPEAT; 5.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS01204; REL_1; 1.
DR PROSITE; PS50254; REL_2; 1.
KW DNA-binding; Transcription regulation; Activator; Nuclear protein;
KW Repeat; ANK repeat; Phosphorylation; Alternative splicing;
KW Disease mutation; Proto-oncogene; 3D-structure.
FT DOMAIN 38 343 REL-LIKE (RHD).
FT REPEAT 486 518 ANK 1.
FT REPEAT 525 554 ANK 2.
FT REPEAT 558 590 ANK 3.
FT REPEAT 598 627 ANK 4.
FT REPEAT 632 662 ANK 5.
FT REPEAT 666 695 ANK 6.
FT REPEAT 728 757 ANK 7.
FT DOMAIN 337 341 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 350 400 GLY-RICH.
FT VARIANT 666 666 A -> SAS (IN TRUNCATED FORM P80HT).
/FTID-VAR_006909.
MISSING (IN TRUNCATED FORM P80HT).
/FTID-VAR_006910.
K -> E (IN REF. 2).
T -> I (IN REF. 2).
R -> G (IN REF. 2).
EP -> DA (IN REF. 2).
HAAPSTRLLRHG -> QRSARALLDYGVTD (IN REF. 2).
L -> K (IN REF. 2).
T -> TT (IN REF. 2).
E -> K (IN REF. 2).
S -> C (IN REF. 2).
H -> TDLLPAPSPPLPGPPVQRPPLHQLFNTPHPLSWDK (IN REF. 1).
FT SEQUENCE 898 AA; 96752 MW; DFD153BE9E23A42B CRG64;

```

alignment_scores:

```

Quality: 224.50 Length: 192
Ratio: 1.887 Gaps: 5
Percent Similarity: 61.979 Percent Identity: 33.854

```

alignment_block:

US-09-509-775-1 x KBF2_HUMAN

Align seg 1/1 to: KBF2_HUMAN from: 1 to: 898

```

181 GATAAATCCCTGCTACTAGAACTGACGAGCAGCAGCACTGTCATGCA 230
|||||..... |||..... |||..... |||..... |||..... |||.....
515 AspleuglyValValasnLeuthrAshHisLeuHisGlnThrProLeuH 531
231 CTGGGCGTCTGCTGACATGACATACAGAAATTTGTAATTTTGTGCAAC 280
| ||| ..... |||..... |||..... |||..... |||..... |||.....
531 sLeuAlaValIlethrGlyGlnThrSerValSerPheLeuLeuArgV 548
281 TTGAGTGCAGTGAATGATAAACACATGCGAGTGTGCTCTCTCTTCAT 330
|||||..... |||..... |||..... |||..... |||..... |||.....
548 alGlyAlaAspProAlaLeuLeuAspArgHisGlyAspSerAlaMetHis 564
331 ATTGCG.....GCTTCTGTCGGCGGATGATGATGATGATGATGATGAT 374
|||||..... |||..... |||..... |||..... |||..... |||.....
565 LeuAlaLeuArgAlaGlyAlaGlyAlaProGlnLeuLeuArgAlaLeuLe 581
375 GGGAAAAGGTGCTCAAGTGAATGCTGCAATCAA..... 408
| ..... |||..... |||..... |||..... |||..... |||.....

```

```

581 uGlnSerGlyAlaPro.....AlaValProGlnLeuLeuHisMetProA 596
409 .....AATGGCTGTACTCCCTTACATTATGCAGCTTCGAAAAACAGGCAT 453
596 spPheGluGlyLeuTyrProValHisLeuAlaValArgAlaArgSerPro 612
454 GAGATCGCTGCTACTTGTGGAAGCGGGGCTAATCCAGATGCTAAAGGA 503
||| ..... |||..... |||..... |||..... |||..... |||.....
613 GluCysLeuAspLeuValAspSerGlyAlaGluValGluAlaThrG1 629
504 CCATTATGAGGCT...ACAGCAATGCACCGGCGAGCAGCCCAAGGTTACT 550
||| ..... |||..... |||..... |||..... |||..... |||.....
629 uArgGlnGlyArgThrAlaLeuHisLeuAlaThrGluMetGluGluL 646
551 TGAAGATGATT...CATATCCTCTGTACTACAAAGCATCCACAAACATC 597
|| ..... |||..... |||..... |||..... |||..... |||.....
646 euGlyLeuValThrHisLeuValThrLysLeuArgAlaAsnValasnAla 662
598 CAAGACACTGAGGGTAACACTCTCTACATTAGCTGTGTATGAGGAGAG 647
||| ..... |||..... |||..... |||..... |||..... |||.....
663 ArgThrPheAlaGlyAsnThrProLeuHisLeuAlaGlyLeuGlyTy 679
648 AGTGAAGAAGCAAACTGCTGTGTCCTCCCAAGGAGCAAGTATTACATTG 697
||| ..... |||..... |||..... |||..... |||..... |||.....
679 rProThrLeuThrArgLeuLeuLysAlaGlyAlaAspIleHisAlaG 696
598 AGAATAAGAGAGAAAGACACACCCCTG 723
|||||..... |||..... |||..... |||..... |||..... |||.....
696 luAsnGluGluProLeuCysProLeu 704
seq_name: SwissProt_39:ILK_MOUSE
seq_documentation_block:
ID ILK_MOUSE STANDARD; PRT: 452 AA.
AC 055222;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE INTEGRIN-LINKED PROTEIN KINASE (EC 2.7.1.-).
GN ILK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=Heart;
RX MEDLINE=98031580; PubMed=9366252;
RA Li F., Liu J., Mayne R., Wu C.;
RT "Identification and characterization of a mouse protein kinase that is
RL Biochim. Biophys. Acta 1358:215-220(1997).
CC -!- FUNCTION: RECEPTOR-PROXIMAL PROTEIN KINASE REGULATING INTEGRIN-
CC MEDIATED SIGNAL TRANSDUCTION. MAY ACT AS A MEDIATOR OF INSIDE-OUT
CC INTEGRIN SIGNALING. FOCAL ADHESION PROTEIN PART OF THE COMPLEX
CC ILK-PINCH. THIS COMPLEX IS CONSIDERED TO BE ONE OF THE CONVERGENCE
CC POINTS OF INTEGRIN- AND GROWTH FACTOR-SIGNALING PATHWAY. COULD BE
CC IMPLICATED IN MEDIATING CELL ARCHITECTURE, ADHESION TO INTEGRIN
CC SUBSTRATES AND ANCHORAGE-DEPENDENT GROWTH IN EPITHELIAL CELLS.
CC PHOSPHORYLATES BETA-1 AND BETA-3 INTEGRIN SUBUNIT ON SERINE AND
CC THRONINE RESIDUES, BUT ALSO AKT1 AND GSK3B (BY SIMILARITY).
CC -!- SUBUNIT: INTERACTS WITH CYTOPLASMIC DOMAIN OF BETA 1 SUBUNIT OF
CC INTEGRIN. COULD ALSO INTERACTS WITH BETA 2, BETA 3 AND/OR BETA 5
CC SUBUNIT OF INTEGRIN, AND ALSO WITH PINCH (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LUNG, HEART, KIDNEY,
CC LIVER, BRAIN, SPLEEN AND SKELETAL MUSCLE. WEAKLY EXPRESSED IN
CC TESTIS.
CC -!- DOMAIN: A PH-LIKE DOMAIN IS INVOLVED IN PHOSPHATIDYLINOSITOL
CC PHOSPHATE BINDING (BY SIMILARITY).
CC -!- PTM: AUTOPHOSPHORYLATED ON SERINE (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 3 ANK REPEATS.

```


Sequence	Strd	Orig	ZScore	EScore	Len	Documentation	...
sp_human:075832	+	1164.00	1881.43	6.4e-97	226	! Q75832 homo sapiens (human)	.28
sp_rodent:Q9A2X3	+	1106.00	1787.22	1.1e-91	231	! Q9A2X3 rattus norvegicus (rat)	.28
sp_rodent:Q9A2X2	+	1105.00	1785.60	1.4e-91	231	! Q9A2X2 mus musculus (mouse)	.28
sp_human:095533	+	582.50	943.92	2.0e-44	118	! Q95533 homo sapiens (human)	.41
sp_plant:Q9Z079	+	338.00	541.58	2.5e-22	247	! Q9Z079 arabidopsis thaliana (ma	.41
sp_rodent:Q61307	+	312.50	483.62	5.4e-20	1943	! Q61307 mus musculus (mouse)	.41
sp_fungi:Q9P3P6	+	310.00	496.52	8.4e-20	237	! Q9P3P6 neurospora crassa, relat	.41
sp_human:Q13484	+	310.00	484.24	8.8e-20	1088	! Q13484 homo sapiens (human)	.41
sp_human:Q12955	+	310.00	473.03	9.3e-20	4377	! Q12955 homo sapiens (human)	.41
sp_rodent:Q70511	+	306.00	470.67	2.1e-19	2622	! Q70511 rattus norvegicus (rat)	.41
sp_invertebrate:Q9T267	+	305.00	480.46	2.5e-19	636	! Q9T267 caenorhabditis elegans	.41
sp_rodent:Q88521	+	301.00	465.77	5.8e-19	1762	! Q88521 rattus norvegicus (rat)	.41
sp_invertebrate:Q9NCP8	+	297.00	462.65	1.3e-18	1159	! Q9NCP8 drosophila melanogast	.41
sp_invertebrate:Q9V5A2	+	297.00	456.65	1.4e-18	2443	! Q9V5A2 drosophila melanogast	.41
sp_human:Q9HA54	+	295.50	459.52	1.8e-18	1265	! Q9HA54 homo sapiens (human)	.41
sp_human:Q9H2K2	+	293.50	456.93	2.7e-18	1166	! Q9H2K2 homo sapiens (human)	.41
sp_mammal:Q9N180	+	290.50	452.28	5.1e-18	1136	! Q9N180 bos taurus (bovine)	.41
sp_human:Q13768	+	290.00	448.13	5.7e-18	1719	! Q13768 homo sapiens (human)	.41
sp_human:Q99407	+	290.00	447.51	5.8e-18	1856	! Q99407 homo sapiens (human)	.41
sp_invertebrate:Q17344	+	288.50	445.39	7.9e-18	1786	! Q17344 caenorhabditis elegan	.41
sp_invertebrate:Q17487	+	288.50	445.29	7.9e-18	1809	! Q17487 caenorhabditis elegan	.41
sp_invertebrate:Q17488	+	288.50	445.26	7.9e-18	1815	! Q17488 caenorhabditis elegan	.41
sp_invertebrate:Q17486	+	288.50	445.03	7.9e-18	1867	! Q17486 caenorhabditis elegan	.41
sp_invertebrate:Q17489	+	288.50	444.32	7.9e-18	2039	! Q17489 caenorhabditis elegan	.41
sp_invertebrate:Q17343	+	288.50	434.39	8.2e-18	6994	! Q17343 caenorhabditis elegan	.41
sp_rodent:Q61302	+	287.50	443.49	9.7e-18	1848	! Q61302 mus musculus (mouse)	.41
sp_rodent:Q97582	+	286.00	447.39	1.3e-17	843	! Q97582 rattus norvegicus (rat)	.41
sp_rodent:Q61304	+	285.50	444.45	1.4e-17	1098	! Q61304 mus musculus (mouse)	.41
sp_human:Q15084	+	283.50	441.50	2.2e-17	1059	! Q15084 homo sapiens (human)	.41
sp_human:Q95271	+	281.50	436.44	3.3e-17	1327	! Q95271 homo sapiens (human)	.41
sp_invertebrate:Q9XZ37	+	281.00	436.56	3.7e-17	1181	! Q9XZ37 drosophila melanogast	.41
sp_invertebrate:Q9VBP3	+	281.00	436.56	3.7e-17	1181	! Q9VBP3 drosophila melanogast	.41
sp_invertebrate:Q92421	+	276.50	427.08	9.5e-17	1549	! Q92421 drosophila melanogast	.41
sp_invertebrate:Q9V4B1	+	276.50	427.08	9.5e-17	1549	! Q9V4B1 drosophila melanogast	.41
sp_mammal:Q9NV77	+	270.50	422.53	3.2e-16	815	! Q9NV77 swi scrofa (pig)	.130 ka
sp_human:Q14974	+	270.50	420.64	3.3e-16	1030	! Q14974 homo sapiens (human)	.41
sp_human:Q62937	+	269.50	422.63	4.0e-16	658	! Q62937 rattus norvegicus (rat)	.41
sp_rodent:Q10728	+	269.50	419.46	4.0e-16	976	! Q10728 rattus norvegicus (rat)	.41
sp_vertbrate:Q90624	+	268.50	417.94	4.9e-16	963	! Q90624 gallus gallus (chicken)	.41
sp_vertbrate:Q90623	+	268.50	417.61	4.9e-16	1004	! Q90623 gallus gallus (chicken)	.41

[illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

```
247 GCACATACAGAAATTTGAATTTTGTGCAACTTGAGTGGCCAGTGAA 296
|||||
51 GlyHisThrGluIleValGluPheLeuLeuGlnLeuGlyValProValas 67
|||||
297 TCATAAAGCAGATGAGGTGGTCTCTCTTCAATTCATGCGGCTTCTGCTG 346
|||||
67 nAspLysAspAlaGlyTrpSerProLeuHisIleAlaSerAlaG 84
|||||
347 GCCGGGATGAGATTGTAAGCCCTCTCGGAAAGCTCAAGTGAAT 396
|||||
84 lYArgSpGluIleValLysAlaLeuLeuGlyLysGlyAlaGlnValasn 100
|||||
397 GGTGTCAATCAAAATGGTGTACTCTCTTACATTTATGAGCTTCGAAAA 446
|||||
101 AlaValasnGlnAsnGlyCysThrProLeuHisIleAlaSerLysas 117
|||||
447 CAGGCATGAGATCGGTGTCATGTTACTGGAAGGGGGCTAATCCAGATG 496
|||||
117 nArgHisGluIleAlaValMetLeuLeuGluGlyAlaAlaSerAlaG 134
|||||
497 CTAAGGACCATTTATGAGCTACAGCAATGCGCCGAGCCCAAGGTT 546
|||||
134 lAlLysAspHisTyGluAlaThrAlaMetHisArgAlaAlaLysGly 150
|||||
547 AACTTGAAGATGATTCATPATCTCTGTPACTACAAAGCATCCACAACAT 596
|||||
151 AsnLeuLysMetIleHisIleLeuLeuTyTrpLysAlaSerThrAsnIl 167
|||||
597 CCAAGACACTGAGGTGATACATCTCTACATTTAGCCTGTGATGAGGAGA 646
|||||
167 eGlnAspThrGluGlyAsnThrProLeuHisIleAlaCysAspGluGlu 184
|||||
647 GAGTGAAGAAGAAAGACCCCTGCTGCTGCCAAGGAGCAAGTATTACATT 696
|||||
184 rgValGluGluAlaLysLeuLeuValSerGlnGlyAlaSerIleTyTrile 200
|||||
697 GAGATAAAGAAAGAAAGACCCCTGCTGCTGCCAAGTGGCCAAAGTGGCTGG 746
|||||
201 GluAsnLysGluGluLysThrProLeuGlnValAlaLysGlyGlyLeuGl 217
|||||
747 TTTAATACCTACAGCAATGGTGAAGGT 774
|||||
217 yLeuIleLeuLysArgMetValGluGly 226
|||||
seq_name: sp_rodent:Q922X3
```

seq_documentation_block:

ID Q922X3 PRELIMINARY; PRT; 231 AA.

AC Q922X3;

DT 01-MAY-1999 (Tremblrel. 10, Created)

DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)

DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)

DE GANKYRIN HOMOLOGUE

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=PLACENTA;

RA Higashitsuji H., Fujita J.;

RT "Cloning of rat gankyrin homologue containing ankyrin repeats.";

RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB022014; BAA36954.1; -

DR HSSP; P42773; 1IHB.

DR InterPro; IPR002110; -

DR Pfam; PF00023; ank; 5.

DR PROSITE; PSS0088; ANK_REPEAT; 5.

DR PROSITE; PSS0297; ANK_REP_REGION; 1.

DR SMART; SM00248; ANK; 1.

SQ SEQUENCE 231 AA; 24985 MW; F5241DC9A816066E CRC64;

alignment_scores:

Quality: 1106.00 Length: 225

Ratio: 4.960 Gaps: 0

Percent Similarity: 99.111 Percent Identity: 94.667

alignment_block:

US-09-509-775-1 x Q922X3 ..

Align seg 1/1 to: Q922X3 from: 1 to: 231

```
97 ATGAGGGGTGTGTCTTAACCTAATGGTGTGCAACTGGCTTACAGGG 146
|||||
1 MetGluGlyCysValSerAsnLeuMetValCysAsnLeuAlaTyAsnGl 17
|||||
147 GAAGCTCGAAGAGTTGAAGGAGATTTCTGGCCGATAAATCCCTGGCTA 196
|||||
17 lYlLeuAspGluLeuLysGluSerIleLeuAlaAspLysSerLeuAla 34
|||||
197 CTAGAACTGACAGGAGCAGCAACTGCATTGCACTGGCATGCTCAGCT 246
|||||
34 hrArgThrAspGlnAspSerArgThrAlaLeuHisTrpAlaCysSerAla 50
|||||
247 GCACATACAGAAATTTGTAATTTTGTGCAACTTGGAGTGGCCAGTGAA 296
|||||
51 GlyHisThrGluIleValGluPheLeuLeuGlnLeuGlyValProValas 67
|||||
297 TGATAAAGACATGCGTGTCTCTCTTCAATTTGCGGCTTCTGCTG 346
|||||
67 nGluLysAspAlaGlyTrpSerProLeuHisIleAlaAlaSerAlaG 84
|||||
347 GCCGGGATGAGATTGTAAGCCCTCTCGGAAAGGTGCTCAAGTGAAT 396
|||||
84 lYArgSpGluIleValLysAlaLeuLeuLysGlyAlaGlnValasn 100
|||||
397 GCTGTCAATCAAAATGGTGTACTCTCTTACATTTATGCAAGTTCGAAAA 446
|||||
101 AlaValasnGlnAsnGlyCysThrAlaLeuHisTyTrpAlaAlaSerLysas 117
|||||
447 CAGCATGAGATCGTGTCTGCTGCTGGAAGGGGGCTAATCCAGATG 496
|||||
117 nArgHisGluIleAlaValMetLeuLeuGluGlyAlaAlaSerProAsp 134
|||||
497 CTAAGGACCATTTATGAGCTACAGCAATGCGCCGAGCCCAAGGTT 546
|||||
134 lAlLysAsnHisTyAspAlaThrAlaMetHisArgAlaAlaLysGly 150
|||||
547 AACTTGAAGATGATTCATPATCTCTGTPACTACAAAGCATCCACAACAT 596
|||||
151 AsnLeuLysMetValHisIleLeuLeuPheTyTrpLysAlaSerThrAsnIl 167
|||||
597 CCAAGACACTGAGGTGATACATCTCTACATTTAGCCTGTGATGAGGAGA 646
|||||
167 eGlnAspThrGluGlyAsnThrProLeuHisIleAlaCysAspGluGlu 184
|||||
647 GAGTGAAGAAGAAAGACCCCTGCTGCTGCCAAGGAGCAAGTATTACATT 696
|||||
184 rgValGluGluAlaLysLeuLeuValSerGlnGlyAlaSerIleTyTrile 200
|||||
697 GAGATAAAGAAAGAAAGACCCCTGCTGCTGCCAAGTGGCCAAAGTGGCTGG 746
|||||
201 GluAsnLysGluGluLysThrProLeuGlnValAlaLysGlyGlyLeuGl 217
|||||
747 TTTAATACCTACAGCAATGGTGAAGGT 771
|||||
217 yLeuIleLeuLysArgIleAlaGlu 225
|||||
```

seq_name: sp_rodent:Q922X2

seq_documentation_block:

ID Q922X2

AC Q922X2; PRELIMINARY; PRT; 231 AA.

DT 01-MAY-1999 (Tremblrel. 10, Created)

DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE GANKYRIN.
GN PSMD10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Higashitsuji H., Fujita J.;
RT "Cloning of mouse gankyrin containing ankryrin repeats.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022022; BAA36969.1; -;
DR HSSP; P42773; 1IIB.
DR MGD; MGI:185898; Psmd10.
DR InterPro; IPR002110; -;
DR Pfam; PF00023; ank; 5.
DR PROSITE; PS50088; ANK_REPEAT; 5.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 231 AA; 25115 MW; 5A3DB7027B0694E6 CRC64;

alignment_scores:
Quality: 1105.00 Length: 225
Ratio: 4.955 Gaps: 0
Percent Similarity: 99.111 Percent Identity: 93.333

alignment_block:
US-09-509-775-1 x Q922X2 ..

Align seg 1/1 to: Q922X2 from: 1 to: 231

97 ATGAGGGGTGTGTCTAACCCTAACTGGTCTGCAACCTGGCTACACGG 146
|||||
1 MetGluGlyCysValSerAsnIleMetIleCysAsnLeuAlaTyrSerG 17
147 GAAGCTGGAAGCTGAAGGAGAGTATTCGGCGGATAAATCCCTGGCTA 196
|||||
17 YLYLeuAlaSpLuleuLysGluArgIleLeuAlaAspLysSerLeuAla 34
197 CTAGAACTGACCGAGCAGAGCACTGCATTGCACTGGCGATGCTCAGCT 246
|||||
34 hrArgThrAspGlnAspSerArgThrAlaLeuHisTrpAlaCysSerAla 50
247 GGACATACAGAAATGTGTGAATTTTCTTCAACTGGAGTGGCGAGTGAA 296
|||||
51 GlyHisThrGluIleValGluPheLeuLeuGlnLeuGlyValProVala 67
297 TGATAAGAGCAGTGCAGGTGGTCTCTCTTCATATTTGGCGGCTCTCTG 346
|||||
67 nAspLysAspAlaGlyTrpSerProLeuHisIleAlaAlaSerAla 84
347 GCCGGGATGAGATTGTAAGACCTTCTGGAAAGTGTCAAGTCAAT 396
|||||
84 lyArgAspGluIleValLysAlaLeuLeuValLysGlyAlaHisValasn 100
397 GCTGTCAATCAAAATGGCTCTACTCCCTTACATTTATGACGCTTCGAAAA 446
:::|||||
101 SerValAsnGlnAsnGlyCysThrProLeuHisTyrAlaAlaSerLysas 117
447 CAGGCATGAGATCGCTCTCATGTTACTTGGAGCGGGCTTAATCCAGATG 496
|||||
117 nArgHisGluIleSerValMetLeuLeuGluGlyAlaAsnProAspA 134
497 CTAAGGACCATTTATGAGCTTACAGCAATTCACCGGCGCAGCAGCAAGGT 546
|||||
134 lalysAspHisTyrAspAlaThrAlamethHisArgAlaAlaAlaLysGly 150
547 AACTTGAAGATGATTTCATCTCTTCTGTACTACAAAGCATCCCAACAT 596
|||||

151 AsnLeuLysMetValHisIleLeuLeuPheTyrLysAlaSerThrAsnIle 167
597 CCAAGACACTGAGGGTAACACTCTCTACACTTACCTGTGATGAGGAGA 646
|||||
167 eGlnAspThrGluGlyAsnThrProLeuHisLeuAlaCysAspGluAla 184
647 GAGTGGGAAGCAAACTGCTGGTGTCCTCCAAAGGAGCAAGTATTTCATT 696
|||||
184 rgValGluGluAlaLysPheLeuValThrGlnGlyAlaSerIleTyrIle 200
697 GAGATAAAGAAAGAACACACCCCTGCAAGTGGCCCAAGGTGCCCTGGG 746
|||||
201 GluAsnLysGluGluLysThrProLeuGlnValAlaLysGlyGlyLeuG 217
747 TTTAATACTCAAGAGAATGGTGGAA 771
|||||
217 yLeuIleLeuLysArgLeuAlaGlu 225
seq_name: sp_human:O95533
seq_documentation_block:
ID O95533 PRELIMINARY;
AC O95533; PRT; 118 AA.
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE DJ889N15.2.2 (26S PROTEASOME SUBUNIT P28 (ANKYRIN REPEAT PROTEIN))
DE (PUTATIVE PARTIAL ISOFORM 2) (FRAGMENT).
GN DJ889N15.2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Graham D.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL031177; CAA20118.1; -;
DR HSSP; O13625; 1YCS.
DR InterPro; IPR002110; -;
DR Pfam; PF00023; ank; 2.
DR PROSITE; PS50088; ANK_REPEAT; 2.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR SMART; SM00248; ANK; 1.
FT NON_TER 118
SQ SEQUENCE 118 AA; 12472 MW; F0475F4C545D15A CRC64;

alignment_scores:

Quality: 582.50 Length: 151
Ratio: 4.936 Gaps: 1
Percent Similarity: 78.146 Percent Identity: 78.146

alignment_block:

US-09-509-775-1 x O95533 ..

Align seg 1/1 to: O95533 from: 1 to: 118

97 ATGAGGGGTGTGTCTAACCCTAACTGGTGTGCAACCTGGCTACACGG 146
|||||
1 MetGluGlyCysValSerAsnLeuMetValCysAsnLeuAlaTyrSerG 17
147 GAAGCTGGAAGCTTGAAGGAGAGTATTCGGCGGATAAATCCCTGGCTA 196
|||||
17 YLYLeuGluGluLeuLysGluSerIleLeuAlaAspLysSerLeuAla 34
197 CTAGAACTGACCGAGCAGCAGCAACTGCATTGCACTGGCGATGCTCAGCT 246
|||||
34 hrArgThrAspGln..... 38
247 GGACATACAGAAATGTGTGAATTTTGTGCAACTGGAGTGGCGAGTGAA 296
38

```

297 TCATAAGACGATGAGTGGTGTCTCTCTTCATATTCGGCTTCTGCTG 346
|||||
39 .....AlaGlyTTPSerProLeuHisIleAlaAlaSerAlaG 51
347 GCCGGGATGAGATGCTAAAGCCCTTCGGGAAGAGTGTCAAGTGAAT 396
|||||
51 lYArgAspGluIleValLysAlaLeuLeuGlyLysGlyAlaGlnValAsn 67
397 GGTGTCATCAAAATGGCTGTACTCCCTTACATATTGACGCTTCGAAAAA 446
|||||
58 AlaValAsnGlnAsnGlyCysThrProLeuHisIleAlaAlaSerLysAs 84
447 CAGGATGAGATGCTGTCTCTTACTTGAAGCGGGGCTTAATCCAGATG 496
|||||
84 nArgHisGluIleAlaValMetLeuLeuGluGlyAlaAsnProAspA 101
497 CTAAGCATTATGAGGCTACAGCAATGCACCGGGGAGCCAGCCCAAGGT 546
|||||
101 lALysAspHisTyrGluAlaThrAlaMetHisArgAlaAlaAlaLysGly 117
547 AAC 549
|||
118 Asn 118
seq_name: sp_plant:Q9ZQ79
seq_documentation_block:
ID Q9ZQ79 PRELIMINARY; PRT; 247 AA.
AC Q9ZQ79;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE PUTATIVE ANKYRIN.
GN T4W8.14.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.,
RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
RA Fraser C.M., Venter J.C.;
RT "Arabidopsis thaliana chromosome II BAC T4W8 genomic sequence.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006284; AADI7433.1; -
DR HSSP; P42773; 1IHB.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank; 5.
DR PROSITE; PS50088; ANK_REPEAT; 5.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 247 AA; 26723 MW; 452884EBD70EFABE CRC64;

```

```

alignment_scores:
  Quality: 338.00      Length: 219
  Ratio: 2.315        Gaps: 6
Percent Similarity: 66.667 Percent Identity: 41.096

```

```
alignment_block:
```

```
US-09-509-775-1 x Q9ZQ79 ..
```

```
Align seg 1/1 to: Q9ZQ79 from: 1 to: 247
```

```

160 TTGAAGGAGAGTATTCGGCCGATAAATCCCTGGCTACTAGAGTACCA 209
|||||
31 LeuSerGluGluGlnLeuSer...LysSerLeuAsnPheArg...AsnG 45

```

```

210 GGACAGCAGAACTCCATTTGCATGGGCGATGCTCAGCTGGACATACAGAA 259
|||||
45 uaspGlyArgSerLeuLeuHisValAlaAlaSerPheGlyHisSerGlnI 62
260 TTGTTGAATTTTGTG.....CAACTTGGAGTCCCAAGTGAATCAT 300
|||||
62 leValLysLeuLeuSerSerAspGluAlaLysThrValIleAsnSer 78
301 AAACAGCATGAGTGTGTCTCTCTTCAATATTCGGCTTCTGCTGCGCG 350
|||||
79 LysaspaspGluGlyTTPAlaProLeuHisSerAlaAlaSerIleGlyAs 95
351 GGATGAGATGCTAAAGCCCTTCTGGGAAAGGTGCTCAAGTGAATGCTG 400
: |||||
95 nAlaGluLeuValGluLeuLeuThrArgGlyValaAspValaAsnAla 112
401 TCAATCAAAATGGCTGTACTCCCTTACATATTGACGCTTCGAAAAACAGG 450
|||||
112 ysAsnAsnGlyGlyArgThrAlaLeuHisTyrAlaAlaSerLysGlyArg 128
451 CATGAGATGCTGTCTTACTTGAAGCGGGGCTAATCCAGATGCTAA 500
|||||
129 LeuGluIleAlaGlnLeuLeuLeuThrHisGlyAlaLysIleAsnIleTh 145
501 GGACATTATGAGGCTACAGCAATGCACCGGGGAGCCAGCCAGGTAAC 550
|||
145 rAspLysValGlyCysThrProLeuHisArgAlaAlaSerValGlyLysL 162
551 TGAAGATGATTCAATATCCTTCTGTACTACAAGCATCCACAACATCCAA 600
|||||
162 euGluValCysGluPheLeuIleGluGluGlyAlaGluIleAspAlaThr 178
601 GACACTGAGGTAACACTCCCTCTA.....CACTTAGCTGTGTAGGGA 644
|||
179 AspLysMetGlyGlnThrAlaLeuMetHisSerValIleCysaspAspLys 195
645 GAGAGTG.....GAAGAAGCAAACTGCTGTGTCCTCCCAAG 679
|||||
195 sGlnLeuLysValSerMetAspGlnValAlaPheLeuLeuIleArgHisG 212
680 GAGCAAGTATTACATTCGAGTAATAAGAAAGAACACACCCCTGCAAGTG 729
|||||
212 lYAlaAspValaspValGluaspLysGlyGlyTyrThrValLeuGlyArg 228
730 GCCAAGTGCGCTGGGTTTAATATCT.....AAGAGAATGGT 767
|||
229 AlathrasnGluPheArgProAlaLeuIleAspAlaAlaLysAlaMetLe 245
768 GGAAGGT 774
|||||
245 uGluGly 247
seq_name: sp_rodent:Q61307
seq_documentation_block:
ID Q61307 PRELIMINARY; PRT; 1943 AA.
AC Q61307; Q61306; Q61308; Q61310; Q61310; Q08866; Q08867;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ANKYRIN 3 (ANKYRIN G) (EPITHELIAL ANKYRIN) (ANKYRIN-3).
GN ANK3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1-6).
RC STRAIN=C57BL/6J; TISSUE=KIDNEY;
RX MEDLINE=95340633; PubMed=7615634;
RA Peters L.L., John K.M., Lu F.M., Elcher E.M., Higgins A., Vialamas M.,
RA Turtzo L.C., Olsuka A.J., Lux S.E.;
RT "Ank3 (epithelial ankyrin), a widely distributed new member of the

```


683 oLeuHisLeuAlaAlaGlnGluAspArgValAsnValAlaGluValLeuV 700
 671 TGTCCAGGAGCAAGTAGTATTACATTGAGATAAAGAAAGACACCC 720
 |||:||||||| ||| :||||||| ||| :||||||| ||| :|||||||
 700 AlAsnGlnGluAlaHisValAspAlaGlnThrLysMetGlyTyrThrPro 716

721 CTGCAAGTG 729
 |||:|||||
 717 LeuHisVal 719

seq_name: sp_fungi:Q9P3P6

seq_documentation_block:

ID Q9P3P6 PRELIMINARY; PRT; 237 AA.
 AC Q9P3P6;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE RELATED TO 26S PROTEASOME SUBUNIT P28.
 GN B7F18.30.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL389891; CAB97304.1; -
 DR InterPro: IPR002110; -
 DR Pfam: PF00023; ank; 6.
 DR PROSITE; PS50088; ANK_REPEAT; 3.
 DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
 DR SMART; SM00248; ANK; 1.
 KW Proteasome.
 SQ SEQUENCE 237 AA; 25995 MW; 6E2B6EFC69D909B1 CRC64;

alignment_scores:

Quality: 310.00 Length: 226
 Ratio: 2.279 Gaps: 4
 Percent Similarity: 60.177 Percent Identity: 34.513

alignment_block:

US-09-509-775-1 x Q9P3P6 ..

Align seg 1/1 to: Q9P3P6 from: 1 to: 237

136 GCTACAGCGGAAGCTGAAGAGTTGAAGAGATATTCTTGGCGATAA 185
 ||| :||||||| ||| :||||||| ||| :||||||| ||| :|||||||
 12 AlaArgAspGlyLysAlaSerIleValGluSerLeuLeuAsnAlaAsnPr 28
 186 ATCCCTGGCTACTAGCACTGACCAGCAGACAGCAACTGCATTGCACTGGG 235
 :||||||| ||| :||||||| ||| :||||||| ||| :|||||||
 28 OlysLeuAlaGlnArgLysAspAspGlyArgLeuProIleHisTrpA 45
 236 CATGCTCAGCTGGACATACAGAAATTTGTAATTTTGTG...CAACTT 282
 ||||| ||| :||||||| ||| :||||||| ||| :||||||| ||| :|||||||
 45 IeCysSerTyrAsnArgLysGluValValGluLeuLeuValAsnGlnLys 61
 283 GGAGTCCAGTCAGTCAATCAAAAGACAGCAGCTTGGTCTCTTCATAT 332
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 62 GlyPheAspProAspValGluAspMetGlyTyrThrProPheMetI 78
 333 TCGCGCTCTCTGCT...GGCCGGGATGAGATTCTAAAGCCCTTCCTGGAA 379
 ||:||||||| ||| :||||||| ||| :||||||| ||| :|||||||
 78 eSerAlaSerValLysAspSerAspAlaIleIleAspLeuLeuSera 95

380 AAGGTGCTCAAGTGAATGCTCTCAATCAAAATGGCTGTACTCCCTTACAT 429
 :||||||| ||| :||||||| ||| :||||||| ||| :|||||||
 95 rgGlyAlaAspIleAsnGlnThrAsnHisGlnSerGlnThrAlaLeuHis 111
 430 TATCAGCTTCGAAAACACAGCAGCATCAGATCGCTGTCATGTTACTGGAAGG 479
 :||| :||||||| ||| :||||||| ||| :||||||| ||| :|||||||
 112 PheIleAlaSerLysAsnAsnIleAspLeuAlaArgLysLeuLeuSerPr 128
 480 CGGGCTAATCCAGAT.....GCTAAGGACCATTTATGAGG 514
 :||| :|||||||
 128 oAspMetLysProLysProAlaSerValArgValLysAspLysArgGlyG 145
 515 CTACAGCAATGACGGGCGAGCAGCAAGGTAAGTCAAGTGAAGTATTCAT 564
 :||| :||||||| ||| :||||||| ||| :||||||| ||| :|||||||
 145 LntyrProLeuHisArgAlaAlaIleGlySerValProMetIleAsn 161
 565 ATCCTTCTGTAACAAGCATCCACAACATCCCAAGACACTGAGGGTAA 614
 :||||||| :||||||| ||| :||||||| ||| :||||||| ||| :|||||||
 162 LeuLeuLeuGlnHisLysSerProIleAsnAlaSerAspAsnAlaGlyTyr 178
 615 CACTCTCTACATTAAGCTGTGATGAGGAGAGAGTGAAGAAGCAAAAC 664
 :||||||| ||| :||||||| ||| :||||||| ||| :|||||||
 178 rThrProLeuHisAlaValAlaGluGlyHisGlyHisAlaAlaVala 195
 665 TCGTGTCTCCCAAGGAGCAAGTATT..... 690
 :||||||| ||| :||||||| ||| :||||||| ||| :|||||||
 195 laLeuLeuLysAlaGlyAlaGluThrAspLysLysAspMetAspGlyTyr 211
 691TACATTGAGAA 701
 212 LeuAlaLeuAspLeuAlaProAspLysAspValArgArgPheIleGluLys 228
 702 TAAAGAAGAAAAGACACCCCTGCAAGTG 729
 :||| :||||||| ||| :||||||| ||| :||||||| ||| :|||||||
 228 sGluAlaGluLysGluGlyIleGluLeu 237

seq_name: sp_human:Q13484

seq_documentation_block:

ID Q13484 PRELIMINARY; PRT; 1088 AA.
 AC Q13484;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE ANKYRIN G119.
 GN ANK3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RX MEDLINE=96234670; PubMed=8666667;
 RA Kashgarian M., Morrow J.S., Ardito T., Stabach P.R., Mann A.S.,
 RA Devatajani P.;
 RT "Identification of a small cytoplasmic ankyrin (AnkG119) in the kidney
 and muscle that binds beta I sigma spectrin and associates with the
 Golgi apparatus."
 RL J. Cell Biol. 133:819-830(1996).
 DR EMBL; U43965; AAB08437.1; -
 DR HSSP; P55273; 1BD8.
 DR InterPro; IPR000906; -
 DR InterPro; IPR002110; -
 DR Pfam; PF00023; ank; 12.
 DR Pfam; PF00791; ZUS; 1.
 DR PROSITE; PS50088; ANK_REPEAT; 12.
 DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
 DR SMART; SM00218; ZUS; 1.
 SQ SEQUENCE 1088 AA; 119427 MW; 769C88D40A78DE86 CRC64;

alignment_scores:

*2.2 Mb of contiguous nucleotide sequence from chromosome III of C.

```

RT RT elegans.
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Graves T., Sutterer C., Ozersky P.;
RT "The sequence of C. elegans cosmid F40G9."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF099919; AAC68798.1; -.
DR HSSP; Q00421; IAWC.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank; 5.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 636 AA; 73229 MW; AB55162AF5D5B1C0 CRC64;

```

alignment_scores:
 Quality: 305.00 Length: 294
 Ratio: 2.061 Gaps: 6
 Percent Similarity: 50.340 Percent Identity: 29.252

alignment_block:

```

US-09-509-775-1 x Q9T267 ..
Align seg 1/1 to: Q9T267 from: 1 to: 636

109 GTGCTTAACCTAATGCTGTCGCAAC...CTGGCTACAGCGGAAGCTGGA 155
   :: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
335 ILeuLeuHisLeuMetCysIleLeuHisTyrPheThrIlyAsnVa 351
156 AGAGTTGAGGAGAGATGTTCTGCGCCGATAAATCCCTGGCTACTAGAACNG 205
   ||| ||::: ::::: ::::: ::::: ::::: ::::: :::::
351 lGluAlaLysArgLeuLeuThrArgTyrProLysLeuValGlyTyrThra 368
206 ACCAGGACAGCAGCACTGCTACCTGCGCAGTCAGCTGCAGTCGACATACA 255
   ||::: ::::: ::::: ::::: ::::: ::::: :::::
368 spAspSerGlyArgSerThrIleHisPheAlaIaValGlyGlySerLeu 384
256 GAAATTGTTGAATTTTGTGCAACTGGAGTGGAGTGCCAGTGAATGAATAAGA 305
   ::::: ::::: ::::: ::::: ::::: ::::: :::::
385 ProLeuLeuGlnPheAlaIle.....LeuAsnAspProG1 396
306 CGATGCA..... 312
396 uMetAlaHisLysThrAspValArgValGlyMetTrpAsnLeuThrg 413
312 ..... 312
413 luLysIleGluAsnLeuPhePheAsnArgLysIleCysAspPheArgLys 429
312 ..... 312
430 AsnTyrSerLysAsnThrLysAsnProHisPheSerGluPhePheG1 446
312 ..... 312
446 uThrIleAspPheGluHisGluAsnSerGlnAsnAlaLeuLysLysPheP 463
312 ..... 312
463 heSerLeuLysIleAspIlePheAspLeuAsnCysLeuIleLeuProLeu 479
313 GGTGTGCTCTCTTCATATTGCGGCTTCTGCTGCGCGGATGAGATTGT 362
   ||||| ::::: ::::: ::::: ::::: ::::: ::::: :::::
480 GlyTrpThrProLeuMetIleAlaSerSerAlaGlyArgValAspValVa 496

```

```

363 AAAAGCCCTTCTGGGA...AAAGGTGCTCAAGTGAATGCTGTCATCAATA 409
   |::: ||||| ::::: ::::: ::::: ::::: ::::: :::::
496 lArgTyrLeuLeuThrLeuProAspValAspValLysHisThrAsnSera 513
410 ATGGCTGTACTCCCTTACATTATGACGTTCGAAACACAGGCATGAGATC 459
   || ||| ||||| ::::: ::::: ::::: ::::: :::::
513 snLysGlnThrSerLeuHisTyrAlaCysSerLysAsnHisValGluIle 529
460 GCTGTCATGTTACTGGAAGCGGGGCTAAT...CCAGATGCTAAGGACCA 506
   :: ::::: ::::: ||| ::::: |||
530 ValLysLeuLeuIleGluAlaAspProAsnIleAlaAsnLeuProAspLy 546
507 TTATGAGCTACAGCAATGCACGGGCAGCAGCAAGGTAAGTCACTGAAGA 556
   ::::: ||||| ::::: ::::: ::::: ::::: :::::
546 spHeGlyAlaThrAlaLeuHisArgAlaAlaSerArgGlyAsnAspValI 563
557 TGATTCATATCCTTCTGTACTAC...AAAGCATCCACAACATCCAGAC 603
   ::::: ||||| ::::: ::::: ::::: ::::: :::::
563 leValArgAlaLeuValSerThrGlyLysCysSerLeuAspArgGlnAsp 579
604 ACTGAGGTTAACTCCTCTACACTTAGCTGTGATGAGGAGAGAGTGA 653
   ||||| ::::: ||||| ::::: ::::: ::::: :::::
580 GlyGluGlyAsnThrAlaLeuHisLeuAlaCysAspGluAsnArgGlyAs 596
654 AGAAGCAAAAACCTGCTGTCGCCAAGGCAAGTATTTACATTGAGAATA 703
   ::::: ||||| ::::: ::::: ::::: ::::: :::::
596 pValAlaIleLeuValAsnArgGlyAlaAspMetLysMetLeuAsnL 613
704 AAGAGAAAAGCACCCCTGCAAGTGCGCCAAA 735
   ||||| ::::: ||||| :::::
613 ysgLulysGlnThrProLeuGluMetLeuLys 623

```

seq_name: sp_rodent:088521

seq_identification_block:

```

ID O88521 PRELIMINARY; PRT: 1762 AA.
AC O88521;
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE 190 KDA ANKYRIN ISOFORM.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98395113; PubMed=9727010;
RA Thevananthar S., Kolli A.H., Devarajan P.;
RT "Identification of a novel ankyrin isoform (AnkG190) in kidney and
RT lung that associates with the plasma membrane and binds alpha-Na, K-
RT ATPase."
RL J. Biol. Chem. 273:23952-23958(1998).
DR EMBL; AF069525; AAC34809.1; -.
DR HSSP; P55273; 1B18.
DR InterPro; IPR000169; -.
DR InterPro; IPR000488; -.
DR InterPro; IPR000906; -.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; UNKNOWN_1.
DR SMART; SM00005; DEATH; 1.
SQ SEQUENCE 1762 AA; 191898 MW; 9023280086A7BF4E CRC64;

```

alignment_scores:
 Quality: 301.00 Length: 192

```

Ratio: 2.263          Gaps: 0
Percent Similarity: 69.271  Percent Identity: 35.938

alignment_block:
US-09-509-775-1 x 088521  ..

Align seg 1/1 to: 088521 from: 1 to: 1762

154 GAAGAGTTGAAGGAGAGATATTCTTGGCCGATAAATCCTCGCTACTAGAAC 203
|||||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
539 GluaspValaLaAlaPheLeuLeuAspHisGlyAlaPheLeuSerIleTh 555
|||||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
204 TGACCAGGACAGACAGACATTGCTGGCGATCCTCAGCTGGACATA 253
|||||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
555 rThrLysGlyPheThrProLeuHisValaLaAlaLysTyrglyLysL 572
|||||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
254 CAGAAATGTTGAATTTTTGTTGCAACTTGGAGTGCAGTGAATGATAAA 303
|||||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
572 euGluValaLaSerLeuLeuGlnLysSerAlaSerProAspAlaLa 588
|||||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
304 GACGATGCAGGTTGGTCTCTCTTCATTATTCGGCGTCTCGTGGCCGGA 353
|||||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
589 GlyLysSerGlyLeuThrProLeuHisValaLaAlaHisTyraspAsn 605
|||||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
354 TCAGATTGTTAAAGCCCTTCGTGGAAAGGTCTCAAGTGAATGCTGTCA 403
|||||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
605 nLysValaLaLeuLeuLeuaspGlnGlyAlaSerProHisAlaLa 622
|||||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
404 ATCAAAATGGCTGTACTCCCTTACATTATTCAGCTTCGCAAAACAGGCAT 453
|||||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
622 laLysasnGlyTyThrProLeuHisIleAlaLaLysLysAsnGlnMet 638
|||||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
454 GAGATCGCTGTCATGTTACTTGAAGGGGGGCTTAATCCAGATGCTAAGA 503
|||||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
639 AspileAlaThrSerLeuLeuGlyTyrglyAlaAspAlaAsnAlaValTh 655
|||||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
504 CCATTATGAGGCTACAGCAATGCACCGGGCAGCAGCCAAAGGGTAATCGA 553
|||||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
655 rArgGlnGlyIleAlaSerValHisLeuAlaLaGlnGlyHisVala 672
|||||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
554 AGATGATTCAATATCTCTGCTACTACAAAGCTCCACAACATCCCAAGAC 603
|||||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
672 spMetValSerLeuLeuLeuSerArgasnAlaAsnValAsnLeuSerAsn 688
|||||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
604 ACTGAGGGTAACACTCTCTACACTTACGCTGTGATGAGGAGAGATGGA 653
|||||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
689 LysArgGlyLeuAsnProLeuHisLeuGlyGlyGlnGluAspArgValas 705
|||||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
654 AGAAGCAAAACTGCTGGTGTCCCAAGGAGCAAGTATTTACATTGACAATA 703
|||||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
705 nValAlaGluValLeuValAsnGlnGlyAlaHisValaAspAlaGlnThrL 722
|||||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
704 AAGAAGAAAGACCCCTCGCAAGTG 729
|||||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
722 ysMetGlyTyThrProLeuHisVal 730

seq_name: sp_invertebrate:Q9NCP8

seq_documentation_block:
ID Q9NCP8 PRELIMINARY; PRT; 1159 AA.
AC Q9NCP8
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ANKYRIN 2.
GN ANK2 OR CG7462.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Insecta;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7221;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=20304926; PubMed=10844021;
RA Bouley M., Tian M.-Z., Faisley K., Shen Y.-C., Malhotra J.D.,
RA Hortsch M.;
RT "The L1-type cell adhesion molecule influences the
RT stability of neural ankyrin in the Drosophila embryo but not its
RT axonal localization.";
RL J. Neurosci. 20:4515-4523(2000).
DR EMBL: AF190635; AAF73309.1; -.
DR FlyBase: FBgn0017645; Ank2.
DR InterPro: IPR000906; -.
DR InterPro: IPR002110; -.
DR Pfam: PF00023; ank; 24.
DR Pfam: PF00791; ZUS; 1.
DR PROSITE: PS50088; ANK_REPEAT; 21.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR SMART: SM00248; ANK; 1.
SQ SEQUENCE 1159 AA; 125769 MW; 9ED146C9E483DE58 CRC64;

alignment_scores:
    Quality: 297.00      Length: 239
    Ratio: 2.034         Gaps: 4
Percent Similarity: 61.088 percent Identity: 29.707

alignment_block:
US-09-509-775-1 x Q9NCP8 ..

Align seg:1/1 to: Q9NCP8 from: 1 to: 1159

169 AGTATTCTGGCGCGATATA.....TCCTCGCTACTAGACCTGACCCAGGA 212
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
235 SerLeuLeuGluLysGlyGlyAsnIleGluAlaLysThr...ArgAs 270
213 CAGCAAACTCATTCGACTGGCGATGCTCAGCTGGCAGTACACAAATG 262
|::: ||| |||||:|||||:|||||:|||||:|||||:|||||:
270 pGlyLeuThrProLeuHisCysAlaAlaArgSerGlyHisGluGlnVal 287
263 TTGAATTTTGTTCGAACCTGGAGTGGCTGCAGTGAATGAAGACGATGCA 312
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
287 aLAspMetLeuGluArgGlyAlaProIleSerAlaLysThrLysAsn 303
313 GGTTCGTCCTCTCT..... 324
|||:|||||:
304 GlyLeuAlaProLeuHisMetAlaAlaGlnGlyCyluHisValAspAla 320
324 ..... 324

320 aArgIleLeuLeuTyHisArgAlaProValAspGluValThrValAspT 337
325 .....CTTCATATTCCGCTTCTGCTGGCCGGGATGAGATTGTA 363
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
337 yrLeuThrAlaLeuHisValAlaAlaHisCysGlyHisValArgValAla 353
364 AAGGCCCTTCGGGAAAGGTGCTCAAGTGAATGCTGCTCAATCAAAATGG 413
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
354 LysLeuLeuAspArgAsnAlaAspAlaAsnAlaArgAlaLeuAsnG 370
414 CTGTACTCCCTTACATTATGCACTTCGAAACACAGGCATGAGATCGCTG 463
| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
370 yPheThrProLeuHisIleAlaCysLysLysAsnArgLeuLysValVal 387
464 TCATGTTACTGGAAGCGGGGCTTAATCCAGATGCTAAGGACCATATGAG 513
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
387 luLeuLeuLeuArgHisGlyAlaSerIleSerAlaThrGluSerGly 403
514 GCTACGACCAATGCACCGGCGACGACCAAGGGTAACTTGAAGATGATCA 563
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
404 LeuThrProLeuHisValAlaAlaPheMetGlyCysMetAsnIleVal 420
564 TATCCTCTCTGACTACAAGCATCCACAACATCCAGACACTGAGGGTA 613
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
420 etVrLeuLeuGlnHisAspAlaSerProAspValProThrValArgGly 437

```

seq_name: sp_invertebrate:Q9NCP8

```
seq_documentation_block:
ID Q9NCP8 PRELIMINARY; PRT; 1159 AA.
```

DT	01-OCT-2000 (TReMBLrel. 15, Created)
DT	01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT	01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE	ANKYRIN 2.
GN	ANK2 OR CG7462.
OS	Drosophila melanogaster (Fruit fly).
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Musc
OX	Ephydroidea; Drosophilidae; Drosophila.
	NCBI_TaxID=7227;
	[1]
RN	SEQUENCE FROM N.A.
RP	


```
664 CTGCTGGTCCCAAGGACGAAGTATTACATTGAGATAAAGAGAA 713
      ::::: ::::: ::::: ::::: ::::: :::::
454 IleLeuLeuArgAsnGlyAlaGlnValAspAlaArgGluGlnG 470
      ::::: ::::: ::::: ::::: ::::: :::::
714 GACACCCCTGCAAGTGGCCAAA.....GGTGGCTGGTTTAATCTCA 757
      ::::: ::::: ::::: ::::: ::::: :::::
470 nThrProLeuHisIleAlaSerArgLeuGlyAsnValAspIleValMetL 487
      ::::: ::::: ::::: ::::: ::::: :::::
758 AGAATGGTGAAGGT 774
      ::::: :::::
487 euLeuLeuGlnHisGly 492
      ::::: :::::
seq_name: sp_human:Q9HAS4
seq_documentation_block:
ID Q9HAS4 PRELIMINARY; PRT; 1265 AA.
AC Q9HAS4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE TANKYRASE-RELATED PROTEIN (FRAGMENT).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Monz D.W., Meese E.;
RT "Novel tankyrase-related gene.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF305081; AAG25674.1; -
FT NON_TER 1
SQ SEQUENCE 1265 AA; 136905 MW; 3BFD7DE7AEBB0C38 CRC64;

alignment_scores:
  Quality: 295.50      Length: 256
  Ratio: 1.824        Gaps: 4
Percent Similarity: 63.281 Percent Identity: 31.641

alignment_block:
US-09-509-775-1 x Q9HAS4 ..
Align seg 1/1 to: Q9HAS4 from: 1 to: 1265
31 CGTACCGGAGCGCGGAGCTGAGCGGGCTGCTCGCGGACAAAGTAG 80
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
99 ArgIleMetSerGlyArgArgCysAlaGlyGlyAlaAlaCysAlaSe 115
      l::: ::::: ::::: ::::: ::::: :::::
81 TTGCTGGGACAGCAATGGAGGGGTGTGTCTAACCTAATG...GTCT 127
      l::: ::::: ::::: ::::: ::::: :::::
115 rAlaAlaAlaGluAlaValGluProAlaAlaArgGluLeuPheGluAlaC 132
      l::: ::::: ::::: ::::: ::::: :::::
128 GCAACCTGGCCTACAGCGGGAAGCTGGAAGATTGAAGGAGATTCTCTG 177
      |||: ::||: ::|||: ::|||: ::|||: ::|||:
132 ysArg.....AsnGlyAspValGluArgValLysArgLeuValThr 145
      |||: ::||: ::|||: ::|||: ::|||: ::|||:
178 GCCGATAAATCCCTGGCTACTAGACTACCCAGGACAGCAAGTGCATT 227
      ::|||: ::||: ||| ::||: ||| ::||: |||
146 ProGluLysValAsnSerArgAspThrAlaGlyArgLysSerThrProLe 162
      ::|||: ::||: ||| ::||: ||| ::||: |||
228 GCACGTGGCATGCTCAGCTGACATACAGAAATTTGTAATTTTGTTCG 277
      ||||: ::|||: ::|||: ::|||: ::|||: ::|||:
162 uHisPheAlaAla***PheGlyArgLysAspValValGluIlyrLeuLeuG 179
      || |||: ::|||: ::|||: ::|||: ::|||:
278 AACTTGGAGTGCCAGTGAATGATAAGACGATGCAGTTGGTCTCTCTCT 327
      || |||: ::|||: ::|||: ::|||: ::|||:
179 lnAsnGlyAlaAsnValGlnAlaArgAspGlyGlyLeuIleProLeu 195
      || |||: ::|||: ::|||: ::|||: ::|||:
328 CATATTGGCGCTTCTGCTGGCGGATGAGATTGTAAAGCCCTTCTGGG 377
      ||| |||: ::|||: |||: |||: |||: |||:
196 HisAsnAlaCysSerPheGlyHisAlaGluValValAsnLeuLeuLeuAr 212
```

```
378 AAAAGGTGCTCAAGTGAATGCTGTCAATCAAAATGAGTGTACTCCCTTAC 427
      ||||: ::|||: ::|||: ::|||: ::|||: ::|||:
212 gHisGlyAlaAspProAsnAlaArgAspAsnTipAsnTyThrProLeuH 229
      || ||||: ||||: ||||: ||||: ||||: ||||:
428 ATTATGAGCTTCGAAAACAGAGCATGAGATCGCTGTCTACTGTACTGGAA 477
      || ||||: ||||: ||||: ||||: ||||: ||||:
229 isGluAlaAlaIleLysGlyLysIleAspValCysIleValLeuLeuGln 245
      || ||||: ||||: ||||: ||||: ||||: ||||:
478 GCGGGGCTAATCCAGATGCTAAGGAC..... 504
      ||||: ::|||: ::|||: ::|||: ::|||: ::|||:
246 HisGlyAlaGluProThrIleArgAsnThrAspGlyArgThrAlaLeuAs 262
      ||||: ::|||: ::|||: ::|||: ::|||: ::|||:
505 ..... 517
      ||||: ::|||: ::|||: ::|||: ::|||: ::|||:
262 pLeuAlaAspProSerAlaLysAlaValLeuThrGlyGluTyrlLysLysA 279
      ||||: ::|||: ::|||: ::|||: ::|||: ::|||:
518 CAGCAATGACCCGGGAGCAGCCAGGGTAAC...TTGAGATGATTCAT 564
      ::::: ::::: ||||: ||||: ||||: ||||:
279 spGluLeuLeuGluSerAlaArgSerGlyAsnGluGluLysMetMetAla 295
      ::::: ::::: ||||: ||||: ||||: ||||:
565 ATCCTTCTGTACTACAAAGCATCCACAACATCCAGACACTGAGGGTAA 614
      ::|||: ::|||: ::|||: ::|||: ::|||: ::|||:
296 LeuLeuThrProLeuAsnValAspCysHisAlaSerAspGlyArgLysSe 312
      ||||: ::|||: ::|||: ::|||: ::|||: ::|||:
615 CACTCCTCTACACTTAGCCTGTGTGATGAGGAGAGTGAAGAAGCAAAAC 664
      :|||: ||||: ||||: ||||: ||||: ||||:
312 rThrProLeuHisLeuAlaAlaGlyTyArgAsnArgValLysIleValGlnL 329
      :|||: ||||: ||||: ||||: ||||: ||||:
665 TGCTGGTGTCCCAAGGAGCAAGTATTACATTGAGATAAAGAGAAAG 714
      ||||: ::|||: ::|||: ::|||: ::|||: ::|||:
329 euLeuLeuGlnHisGlyAlaAspValHisAlaLysAspLysGlyAspLeu 345
      ::|||: ::|||: ::|||: ::|||: ::|||: ::|||:
715 ACACCCCTGCAAGTGGCC 732
      ::|||: ||||: ||||: ||||: ||||: ||||:
346 ValProLeuHisAsnAla 351
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2001, 07:45:18 ; Search time 1165.93 Seconds
(without alignments)
6323.901 Million cell updates/sec.

Title: US-09-509-775-1
Perfect score: 780
Sequence: 1 tggtagactctaacgctg.....gaatggtaggtaaac 780

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues
Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1:	gb_est1:
2:	gb_est2:
3:	gb_est3:
4:	gb_est4:
5:	gb_est5:
6:	gb_est6:
7:	gb_est7:
8:	gb_est8:
9:	gb_est9:
10:	gb_est10:
11:	gb_est11:
12:	gb_est12:
13:	gb_est13:
14:	gb_est14:
15:	gb_est15:
16:	gb_est16:
17:	gb_est17:
18:	gb_est18:
19:	gb_est19:
20:	gb_est20:
21:	gb_est21:
22:	gb_est22:
23:	gb_est23:
24:	gb_est24:
25:	gb_est33:
26:	gb_est34:
27:	gb_est35:
28:	gb_est36:
29:	gb_est37:
30:	gb_est38:
31:	gb_est39:
32:	gb_est40:
33:	em_estba:
34:	em_estfun:
35:	em_esthum1:
36:	em_esthum2:
37:	em_esthum3:
38:	em_esthum4:
39:	em_esthum5:
40:	em_esthum6:
41:	em_esthum7:
42:	em_esthum8:
43:	em_esthum9:

44:	em_esthum10:
45:	em_esthum11:
46:	em_esthum12:
47:	em_esthum13:
48:	em_esthum14:
49:	em_esthum15:
50:	em_esthum16:
51:	em_esthum17:
52:	em_esthum18:
53:	em_esthum19:
54:	em_esthum20:
55:	em_esthum21:
56:	em_esthum22:
57:	em_esthum23:
58:	em_esthum24:
59:	em_esthum25:
60:	em_esthum26:
61:	em_esthum27:
62:	em_esthum28:
63:	em_estin1:
64:	em_estin2:
65:	em_estin3:
66:	em_estin4:
67:	em_estin5:
68:	em_estom1:
69:	em_estom2:
70:	em_estov1:
71:	em_estov2:
72:	em_estpl1:
73:	em_estpl2:
74:	em_estpl3:
75:	em_estpl4:
76:	em_estpl5:
77:	em_estpl6:
78:	em_estpl7:
79:	em_estpl8:
80:	em_estpl9:
81:	em_estpl10:
82:	em_estro1:
83:	em_estro2:
84:	em_estro3:
85:	em_estro4:
86:	em_estro5:
87:	em_estro6:
88:	em_estro7:
89:	em_estro8:
90:	em_estro9:
91:	em_estro10:
92:	em_estro11:
93:	em_estro12:
94:	em_estro13:
95:	em_estro14:
96:	em_estro15:
97:	em_estro16:
98:	em_estro17:
99:	em_estro18:
100:	em_estro19:
101:	em_estro20:
102:	gb_est25:
103:	gb_est26:
104:	gb_est27:
105:	gb_est28:
106:	gb_est29:
107:	gb_est30:
108:	gb_est31:
109:	gb_est32:
110:	gb_est41:
111:	gb_est42:
112:	gb_est43:
113:	gb_est44:
114:	gb_est45:
115:	gb_est46:
116:	gb_est47:

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	702	90.0	848	108	AU133088	AU133088 AU133088
2	687	88.1	890	106	AL551457	AL551457 AL551457
3	687	88.1	892	106	AL558577	AL558577 AL558577
4	655.6	84.1	767	154	BG499627	BG499627 602545002
5	646.4	82.9	773	173	BG107004	BG107004 602291246
6	641	82.2	777	154	BG499442	BG499442 602546667
7	620	79.5	867	140	BE783372	BE783372 601471858
8	605.8	77.7	783	140	BE785737	BE785737 601478508
9	604	77.4	634	155	BG568346	BG568346 602587280
10	603.8	77.4	915	192	AK009068	AK009068 Mus muscu
11	601	77.1	601	141	BE904296	BE904296 Mus muscu
12	597.8	76.6	646	136	BE504104	BE504104 601494565
13	593.8	76.1	814	154	BG477395	BG477395 602523604
14	590.6	75.7	829	153	BG426275	BG426275 602492455
15	590.2	75.7	1301	192	AK018233	AK018233 Mus muscu
16	584.8	75.0	854	146	BE211764	BE211764 601812392
17	565.4	72.5	587	140	BE787558	BE787558 601479443
18	564.8	72.4	856	141	BE893858	BE893858 601436221
19	560.8	71.9	957	154	BG494082	BG494082 602542388
20	555.8	71.3	786	140	BE787432	BE787432 601479575
21	554.2	71.1	707	154	BG497031	BG497031 602538951
22	552.2	70.8	960	174	BG175434	BG175434 602337789
23	543.8	69.7	942	3	AA203428	AA203428 zx58604.r
24	540.8	69.3	940	152	BG342503	BG342503 602374061
25	528.8	67.8	647	146	BF212788	BF212788 601814051
26	524.8	67.3	624	136	BE539557	BE539557 601060269
27	513.4	65.8	595	173	BG087666	BG087666 H3141E11
28	510	65.4	802	154	BG494451	BG494451 602523438
29	505.4	64.8	910	144	BF122766	BF122766 601761294
30	504.8	64.7	999	153	BG388982	BG388982 602414741
31	503.8	64.6	1274	146	BF241523	BF241523 601878408
32	503.4	64.5	895	168	BF699858	BF699858 602127304
33	497.6	63.8	774	144	BF119890	BF119890 601757668
34	493	63.2	600	31	AV593821	AV593821 AV593821
35	481	61.7	886	174	BG118585	BG118585 602348329
36	474	60.8	547	144	BF077522	BF077522 227450 MA
37	469.8	60.2	550	173	BG085094	BG085094 H3108E09
38	464.6	59.6	745	146	BF217380	BF217380 601885750
39	457.8	58.7	698	32	AV712330	AV712330 AV712330
40	450	57.7	461	7	AA25037	AA25037 zw06a10.r
41	447.8	57.4	765	154	BG529676	BG529676 602558119
42	436	55.9	818	154	BG530573	BG530573 602560158
43	435.8	55.9	495	1	AA035825	AA035825 mhl7d04.r
44	435.4	55.8	652	119	AW682245	AW682245 EST01027
45	434.6	55.7	1002	146	BF217727	BF217727 601882552

ALIGNMENTS

RESULT	1
LOCUS	AU133088
DEFINITION	AU133088 NT2RP4 Homo sapiens cDNA clone NT2RP4001258 5', mRNA
ACCESSION	AU133088
VERSION	AU133088.1
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
	1 (bases 1 to 848)
	Ota, T., Sugiyama, T., Ishii, S., Suzuki, Y., Saito, K., Yamamoto, J.,
	Nishikawa, T., Nakamura, Y., Nagai, T., Sugano, S., Masuho, Y. and
	Isogai, T.
TITLE	HRI human cDNA project (Ota, T., Sugiyama, T., Ishii, S., Suzuki, Y.,

Saito, K., Yamamoto, J., Nishikawa, T., Nakamura, Y., Nagai, T., Sugano, S., Masuho, Y., Isogai, T.)
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 252-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomics@hri.co.jp
HRI human cDNA project; 5' - & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

FEATURES

source
1..848
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RP4001258"
/clone_lib="NT2RP4"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/note="Vector: pME18SPL3; mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT 243 a 168 c 221 g 213 t 3 others
ORIGIN

Query Match 90.0%; Score 702; DB 108; Length 848;
Best Local Similarity 100.0%; Pred. No. 7.2e-200;
Matches 702; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 79 agttgctggagcagcaaaatggagggtgtgtctaaactaagtgtgcaacctggcc 138
Db 1 AGTGTGGGACAGCGAAATGGAGGGGTGTGTCTTAACCTAATGTCTGCAACCTGGGCC 60
QY 139 tacagcgggaagctggaagagttgaaggagagtattctgcccagataaaatccctggctact 198
Db 61 TACACGGGGAAGCTGGACAGAGTGAAGAGAGATATTCTGCCGATAAATCCCTGGCTACT 120
QY 199 agaactgaccagagcagcaactgcatgctgagctggcctgctcagctggacatacagaa 258
Db 121 AGAAGTGCACGAGACAGCAAGTGCATGTCAGTGGCATGCTCAGCTGGACATACAGAA 180
QY 259 attgttgaattttgtgcaacttggagtgccagtggaatgataaaagacatgcagttgg 318
Db 181 ATGTTGGAATTTTGTGCACTTGGAGTGCAGTGAATGATAAAGACATGCGAGTTGG 240
QY 319 tctcctctcatattgcccgtctctgctggccgggagtgagattgtaaaagccctctcgga 378
Db 241 TCTCTCTTCATATGCGGCTTCTGCTGCCGGGATGAGATTGTAAGCCCTTCTGGGA 300
QY 379 aaagtgctcaagtgaatgctgtcattcaaatcaaatggctgtactcccttaccattatgcagct 438
Db 301 AAAGTGTCTCAAGTGAATGCTGCAATCAAAATGGCTGTACTCCCTTACATATATGCAGCT 360
QY 439 tcgaaaaacagcagatgagatgctgtcattgtaaggaggcggggtggtccatccagatgct 498
Db 361 TCGAAAAACAGCATGAGATGCTGTCTGTTACTGGAAGCGGGGCTAATCCAGATGT 420
QY 499 agggaccattatgaggtacagcaatgaccgggagcagcagcaagggttaacttgaagatg 558
Db 421 AAGGACCATATGAGGCTACAGCAATGCAACCGGCGAGCCCAAGGGTAACTTGAAGATG 480
QY 559 attcatatcttctgtactacaagaagcatccacaacaacatccaaagacactgagggtaacact 618
Db 481 ATTATATCTTCTGTGTACTACAAAGCATCCCAACATCCAAAGACACTGAGGGTAACACT 540
QY 619 cctctacacttgccctgtgatgagagagagtggaagaagcaaaactgctggtgtcccaa 678
Db 541 CCTCTACACTTACCTGTGTGATGAGGAGAGTGAAGAGCAAAACTGTGTGTGTCCTCCAA 600
QY 679 ggagcaagtatttaccattgagaataaagaagaaagacacccctgcaagtgccaaaggt 738

```
|||||
Db 601 GGAGCAAGTATTACATTGAGATAAAGAGAAAGACACCCCTGCAAGTGGCCAAAGGT 660
QY 739 ggcctgggttaataactcaagagaagaatggtggaagggttaaca 780
Db 661 GGCTGGGTTTAATACTCAAGAGAATGGTGAAGGTTAAACA 702

RESULT 2
AL551457 AL551457 890 bp mRNA EST 16-FEB-2001
LOCUS AL551457 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI064YP22 5
DEFINITION prime, mRNA sequence.
ACCESSION AL551457
VERSION AL551457.1 GI:12889423
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 890)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1. .890
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 255 a 177 c 221 g 236 t 1 others
ORIGIN

Query Match 88.1%; Score 687; DB 106; Length 890;
Best Local Similarity 100.0%; Pred. No. 2.4e-195;
Matches 687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 gaaatgagggggtgtgtctaaactaatggtctgaacctgacctcagcggaagctg 153
Db 1 GAAATGAGGGGTGTGTCTAACTAATGGTCTGCAACCTGGCTACAGCGGGAAGCTG 60

QY 154 gaagagtgaaggagatattctggccgataaattccctggctactagaactgaccaggac 213
Db 61 GAAGAGTTGAGGAGAGATATTCTGGCCGATTAATCCCTGGCTACTAGACTGACCAGGAC 120

QY 214 agcagaactgcattgcactgggcatgctcagctggacatacagaaaattgtgaattttg 273
Db 121 AGCAGAACTGCATTGCCTGGCAGTCTCAGCTGGCATACAGAAAATTGTTGAATTTTG 180

QY 274 ttgcaacttgagtgccagtgatgataaagacagatcaggttggtctctctcatatt 333
Db 181 TTGCAACTTGGAGTGCCAGTGAATGATAAAGACAGATGAGTTGGTCTCCTCTTCATATT 240

QY 334 gcggcttctgtgcggggatgagattgttaaaagcccttctggaaaagtgctcaagt 393
Db 241 GCGGCTTCTGCTGCGGGGATGAGATTGTAAGACCCCTTCGGGAAAGGTGCTCAAGTG 300
```

```
QY 394 aatgctgtcaatacaaatggctgtactcccttacattatgcagcttcgaaaaaacagcat 453
Db 301 AATGCTGTCAATCAAAATGCTGTACTCCTTACATTATGACGCTTCGAAAAACAGCAT 360
QY 454 gagatcgctgcatgttactggaaggcggggttaataccagatgctgaagaccattatgag 513
Db 361 GAGATCGCTGTCTATGTTACTTGAAGCGGGGGCTTAATCCAGATGCTAAGGACCAATTATGAG 420
QY 514 gctacagcaatgcacggcgcagcagcagcgaagggttaacttgaagatgattcatatccctctg 573
Db 421 GCTACAGCAATGTCACCGGCGAGCAGCCAGGGTAACTTGAAGATGATTATATCTCTCTG 480
QY 574 tactacaagcatcccaacaacatccaagacactgagggtaaacactcctctctacacttagcc 633
Db 481 TACTACAAAGCATCCACAAACATCCAAGACACTGAGGTAACACTCCTCTACACTTAGCC 540
QY 634 tgtgatgaggagagtggaagaagcaaaactcgtcgtgtcccaagaggcaagatatttac 693
Db 541 TGTGATGAGGAGAGAGTGGAAAGAACAAACTGCTGGTGTCCCAAGGAGCAAGTATTTAC 600
QY 694 attgagaataaagaagaaagacacacccctgcaagtggccaaaggtgacctgggtttaata 753
Db 601 ATTGAGAATAAAGAAGAAGAACACCCCTGCAAGTGCCCAAGGTGGCTGGGTTTAATA 660
QY 754 ctcaagagaatggtggaagggttaaca 780
Db 661 CTCAAGAGAATGGTGAAGGTTAAACA 687

RESULT 3
AL558577 892 bp mRNA EST 16-FEB-2001
LOCUS AL558577 LTI_NFL008_TC2 Homo sapiens cDNA clone CS0DJ006YN18 5
DEFINITION prime, mRNA sequence.
ACCESSION AL558577
VERSION AL558577.1 GI:12903232
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 892)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1. .892
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="LTI_NFL008_TC2"
/sex="male"
/tissue_type="T cells from T cell leukemia"
/notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 255 a 177 c 222 g 238 t
ORIGIN

Query Match 88.1%; Score 687; DB 106; Length 892;
Best Local Similarity 100.0%; Pred. No. 2.4e-195;
```


Matches 687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 94 gaaatgaggggtgtgtctaaacctaattggtctcaacctggtcctacagcgggaagctg 153
Db 1 GAAATGGAGGGGTGTGTCTAAACCTAATGTCTCAACCTGGCCCTACAGCGGGAAGCTG 60
Qy 154 gaagagttgaaggagagatttctggtccgataaatacccttggctactagaactgaccagac 213
Db 61 GAAGAGTTGAAGGAGAGTATTCTGGCCGATAAATCCCTGGCTACTAGAACTGACCAGGAC 120
Qy 214 acaacaactgattcactcgtggcgtcagctgacatacagaataatgtgaattttg 273
Db 121 AGCAGAACTGCATTGCATGGGCATGCTCAGCTGACATACAGAAATTTGTAATTTTG 180
Qy 274 ttgcacttgagtgccagtgatgaataaagacgatgcaggttggtctcctctcatatt 333
Db 181 TTGCACACTTGGAGTCCAGTGAATGATGAAGACGATGAGGTGTGTCCTTCATATT 240
Qy 334 gcggcttctgtggtccggatgagattgtaaaagcccttctgggaaaaggtgctcaagt 393
Db 241 GCGGCTTCTGTGGCGGGATGAGATTGTAAAGCCCTTCTGGGAAAGGTGCTCAAGTG 300
Qy 394 aatgctgtaatacaaaatggctgtactccctacattatgcaagcttcgaaaaacaggcat 453
Db 301 AATGCTGTCAATCAAAATGGCTGTACTCCTTACATTATGCAGCTTCGAAAAACAGGCAT 360
Qy 454 gagatcgctgcatgttactgaagcgggctaatccagatgctaagaccattatag 513
Db 361 GAGATCGCTGTCAATGTTACTGGAAGCGGGCTAATCCAGATGCTTAAGACCATTTATGAG 420
Qy 514 gctacagcaatgcacccggcagcagccaaaggtaacttgaagatgattcatctctctg 573
Db 421 GCTACAGCAATGCACCGGCAGCAGCCCAAGGTAACCTGAAGATGATTCATATCCTCTG 480
Qy 574 tactcaagaatcccaaaaatcccaagacactgagggttaacactcctctacactagcc 633
Db 481 TACTAAAGCATCCACAAACATCCCAAGACACTGAGGGTAACACCTCCCTACACTAGCC 540
Qy 634 tctgatgagagagtggaagaagcaaaactgctggtgtcccaaggagcaagtattac 693
Db 541 TGTGATGAGGAGAGTGGAGAGCAAAACCTGCTGGTGTCCCAAGGAGCAAGTATTAC 600
Qy 694 attgagaataagaagaaaagacacccctgcaagtggccaaagggtggcctgggttaata 753
Db 601 ATTGAGATAAAGAGAAAGACACCCCTGCAGTGGCCAAAGGTGGCTGGGTTAATA 660
Qy 754 ctcaagagaatggtggaaggttaaaca 780
Db 661 CTCAAGAGAATGGTGAAGGTTAAACA 687
```

RESULT 4

```
LOCUS BG499627 767 bp mRNA EST 27-MAR-2001
DEFINITION 602545002F1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:4667448 5',
mRNA sequence.
ACCESSION BG499627
VERSION BG499627.1 GI:13461144
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 767)
NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCFD/DRP
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLCM1476 row: c column: 01

High quality sequence stop: 711.

FEATURES

```
Location/Qualifiers
source
1..767
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4667448"
/clone_lib="NIH_MGC_60"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: prostate; Vector: pDNR-LIB (Clontech);
Site_1: SfII (ggcgctcgcc); Site_2: SfII (ggcattatggcc
); Double-stranded cDNA was prepared from cell line RNA.
5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGCCATTATGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCCGAGCGCCGACATG-dT(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.5 kb (range 0.9-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."
```

BASE COUNT 220 a 150 c 210 g 187 t

Query Match 84.1%; Score 655.6; DB 154; Length 767;

Best Local Similarity 97.7%; Pred. No. 6.1e-186;

Matches 686; Conservative 0; Mismatches 14; Indels 2; Gaps 2;

```
Qy 80 gttgctgggacagcgaatgaggggtgtgtcttaacctaatggtctgcaacctggcct 139
Db 3 GTTGTGGGACAGCGAAATGGAGGGGTGTGTCTTAACCTAATGTGCTGCAACCTGGCCT 62
Qy 140 ecagcgggaagctggaagagttgaaggagagtattctggccgataaaatccctggctacta 199
Db 63 ACAGCGGGAAGCTGGAGAGAGTTGAAGAGAGTATTCTGGCCGATTAATCCCTGGCTACTA 122
Qy 200 gaactgaccagagcagcagaactgcattgcactggcgtgctcagctggacatacagaaa 259
Db 123 GAACCTGACAGGACAGCAGAACTGCAATTCATGCGCATGCTCAGCTGGACATACAGAAA 182
Qy 260 ttgtgaatttttgtgcaacttggagtgccagtgaaatgataaagacgatgcaggttggct 319
Db 183 TTGTTGAATTTTGTGTGCAACTTGGAGTGCCAGTGAATGATAAAGACGATGCGAGTTGGT 242
Qy 320 ctccctctcatattgcggtcttctgctggcgggagtgagattgtaaaagccctctgggaa 379
Db 243 CTCCTCTTCATATTGCGGCTTCTGCTGGCCGGATGAGATTGTAAGAGCCCTTCCTGGGNA 302
Qy 380 aagtgctcaagtgaatgctgctcaatacaaaatggcttactcccttaccattatgcagctt 439
Db 303 AAGGTGCTCAAGTGAATGCTGTCAATCAAAATGGCTGTACTCCCTTACATTATGCGAGTT 362
Qy 440 cgaataacagcagcatgagatcgctgcatgttacttgaaggcggggttaactcagatgcta 499
Db 363 CGAAAAACAGGCATGAGATCGCTGTCTACTGGAAGCGGGGCTAATTCAGATGCTA 422
Qy 500 aggaccattatgagctacagcaatgcaccgggagcagcagcgaagggttaacttgaagatga 559
Db 423 AGGACCATTTAGGCTACAGCAATGCACCGGGCAGCAGCCAGGAGTAACTTGAAGATGA 482
Qy 560 ttcatctcttctgtactacaaaagcatccacaataatccaaagacactgagggtaaacactc 619
Db 483 TTCATATCCTTCTGTACTACAAAGCATCCACAAACATCCAAAGACACTGAGGATCAACTC 542
Qy 620 ctctacacttgctgtgtagagagagagtggaagaagcaaaactgctggtgtcccaag 679
Db 543 CTCACACTTAGCTGTGATGAGGAGAGAGTGGAAGAAGCAAAACACTGCTGGTGT-CCAAG 601
```

```

QY 680 gagcaagtattacattgagaataaagaagaagaaagacacccctgcaagtggcgaagaggtg 739
|||||
Db 602 GAGCAAGTATTACATTGAGATTAGAGAGAAAGAACACACCCCTGCAAGTGGCCARAGGTG 661

QY 740 gccct-ggggttaatactcaagaagaagtgtggaagggttaaaca 780
|||||
Db 662 GCCTGGGGTTTAATACTCAAGAGAATGGGTGGGAAGGTTCAAA 703

RESULT 5
BG107004 773 bp mRNA 30-JAN-2001
DEFINITION 602291246F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4386179 5',
mRNA sequence.
ACCESSION BG107004
VERSION BG107004.1 GI:12600850
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 773)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis Staudt, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM0068 row: k column: 12
High quality sequence stop: 655.
Location/Qualifiers
1. .773
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4386179"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: Lymph; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC library."
BASE COUNT 216 a 157 c 207 g 193 t
ORIGIN

Query Match 82.9%; Score 646.4; DB 173; Length 773;
Best Local Similarity 98.8%; Pred. No. 3.6e-183;
Matches 672; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

QY 102 ggggtgtgtctaacctaatgtgtctgcaacctggcctacagcgggaagctggaagatt 161
|||||
Db 1 GGGGTGTGTCTAACTTAATGTGTCTGCAACCTGGCTACAGCGGGAAGCTGGAAGATT 60

QY 162 gaaggagatattctggccgataaactccctggctactagaactgaccagacagcagaac 221
|||||
Db 61 GAAGGAGAGATTCTGGCCGATAAATCCCTGGCTACTAGAACCTGACGACAGCAGAGAAC 120

QY 222 tgcattgcactgggcatgctcagctggacatacagaaaattgttgaaattttgttgcaact 281
|||||
Db 121 TGCATTGCATGGGCATGCTCAGCTGGACATACAGAAATGTTGTAATTTTGTGCAACT 180

QY 282 tggagtgcagtgaaatgataaaagacagatgcaggttggtctcctctcatattcgcccttc 341
|||||
Db 181 TGGAGTGCCAGTGAATGATAAAGACGATGAGGTGTGCTCTCTTCATATTGCGGCTTC 240

```

```

QY 342 tgctggccgggatgagattgtaaaagccctcttggaagagtgctcaagtgaatgctgt 401
|||||
Db 241 TGCTGGCCGGGATGAGATTGTAAAAAGCCCTTCTGGAAAAGGTGCTCAAGTGAATGCTGT 300

QY 402 caatcaaaatggctgtactcccttaccattatcagcttcgaaaaaacaggcatgagatcgc 461
|||||
Db 301 CAATCAAAATGGCTGTACTCCCTTACATTATGAGCTTCGAAAAACAGGCATGAGATCGC 360

QY 462 tgctatgttactggaagcggggttaactccagatcctaagggaccattatgaggtcacagc 521
|||||
Db 361 TGTCTATGTTACTTGGAAAGCGGGGCTAATCCAGATGCTGAAGACCATATGAGGCTACAGC 420

QY 522 aatgcaccgggcagcagcgaagggttaacttgaagatgattcatatccctctgtactacaa 581
|||||
Db 421 AATGCACCGGCGCAGCAGCAAGGTAACTTGAAGATGATTATATCTCTTCTGTACTACAA 480

QY 582 agcatccacaaacatcccaagacactgaggggttaactcactctctacacttagcctgtgatga 641
|||||
Db 481 AGCATCCCAAAACATCCCAAGACACTGAGGGTAACACTCTCTACACTTAGCCCTGTGATGA 540

QY 642 ggagagatgggaagagc-aaactgctgtgtgtcccaaggaagaattttacattgaga 700
|||||
Db 541 GGAGAGAGTGGGAAGAAAGCAAAACACTGCTGTGTCCTCCCAAGAGCAAGTATTTACATTGAGA 600

QY 701 ataaagaagaaagacacccctgcaagtggccaaagtgccctgggttttaatactacaaga 760
|||||
Db 601 ATAAGAAGAAAAAGACACCCCTGCAAGTGCCCAAGGTGG-CTGGGTTTAAATACTCACGA 659

QY 761 gaatgggtgaagggttaaaaca 780
|||||
Db 660 GAATTGTGCGCGGGTTAACA 679

RESULT 6
BG499442 777 bp mRNA 27-MAR-2001
LOCUS 602546667F1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:4669048 5',
mRNA sequence.
ACCESSION BG499442
VERSION BG499442.1 GI:13460959
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 777)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LECMI480 row: e column: 17
High quality sequence stop: 639.
Location/Qualifiers
1. .777
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4669048"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: prostate; Vector: pDNK-LIB (Clontech);
Site:1: SfiI (ggcgcctcgccc); Site:2: SfiI (ggccattatgccc
); Double-stranded cDNA was prepared from cell line RNA.
5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor

```

sequence: 5'-ATTCTAGAGCGGCGGCGGCACATG-DT(30)BN-3'
(where B = A, C, G and N = A, C, G, or T). Average
insert size 1.5 kb (range 0.9-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library.

BASE COUNT 218 a 162 c 213 g 184 t
ORIGIN

Query Match 82.2%; Score 641; DB 154; Length 777;
Best Local Similarity 97.3%; Pred. No. 1.5e-181;
Matches 684; Conservative 0; Mismatches 15; Indels 4; Gaps 3;

QY 80 gtctggagacagcaaatgaggggtgtgtctcaactaatggtctgaacctggcct 139
|||||
Db 3 GTTGTGGGACACGAAATGAGGGGTGTGTCTAACCTAATGGTCTGCAACCTGGCCT 62
|||||
QY 140 acagcgggaagctggaagagtgaaagagagattcttgcgcgataaatccctggctacta 199
|||||
Db 63 ACAGCGGAAGCTGGAAGAGTTCAAGAGAGATATTCTGGCCGATAAATCCTGGCTACTA 122
|||||
QY 200 gaactgaccagcagcagcaactgcattgcactggcgtgcatgctcagctggacatacagaaa 259
|||||
Db 123 GAACTGACCAGGACAGCAGAACTGCATTGCACCTGGCATCTCAGCTGGACATACAGAAA 182
|||||
QY 260 ttgtgaattttgtgcaacttggagtgccagtggaatgaataaagcagatgcaaggttgt 319
|||||
Db 183 TTGTTGAATTTTGTGCAACTTGGAGTGCCAGTGAATGATAAAGACGATGCGAGTTGGT 242
|||||
QY 320 ctctcttcataattgcgctcttgcgtggcggagtgagattgttaaagccctctgggaa 379
|||||
Db 243 CTCTCTTCATATTGGCGCTTCTGCTGGCCGGGATGAGATTGTAAAGCCCTCTGGGAA 302
|||||
QY 380 aagtgctcaagtgaatgtctgtcaatacaaaatggctgtactccctacattatgcagctt 439
|||||
Db 303 AAGTGCTCAAGTGAATGTCTCAATCAAAATGGCTGTACTCCCTTACATTTATGCAGCTT 362
|||||
QY 440 cgaataacagcagatgagctgtctgttactggaagcggggttaactcaagatgcta 499
|||||
Db 363 CGAAACAGGCATGATGATGCTGTCTGCTGTCTGCTGTCTGCTGTCTGCTGTCTGCT 422
|||||
QY 500 aggaccattataggctacagcaatgcacccggcagcagcagcagcagcagcagcagcagc 559
|||||
Db 423 AGGACCATTTAGAGCTACAGCAATGCACCGGCAGCAGCAGCAGCAGCAGCAGCAGCAG 482
|||||
QY 560 ttcatatctctgtactacaaagcattccacaaacatccaagacactgagggtaacactc 619
|||||
Db 483 TTCTATCTCTTCTGTACTACAAAGCATCCACAAACATCCAAAGACACTGAGGGTAACACTC 542
|||||
QY 620 ctctacacttagctgtgagagagagtggaagcagcagcagcagcagcagcagcagcagc 678
|||||
Db 543 CTCTACACTTAGCCTGTGATGAGGAGAGAGTGAAGAGAGCAAACTGCTGGTGTCCCAA 602
|||||
QY 679 ggagcaagtatattacattgagaataaagaagaagacacccctgcaagtggcacaaggt 738
|||||
Db 603 GGGAGCAAGTATTACATTGAGATATAAGAGAAA--GACACCTTGCAGTGGCCCAAGGGT 660
|||||
QY 739 ggcctgggttttaatact-caagaagaatggtggaaggttaaaaa 780
|||||
Db 661 GGCTGGGTTTAACTCCACGAGACTGGTGAAGGTAAACA 703
|||||

RESULT 7
BE783372 BE783372 867 bp mRNA EST 20-OCT-2000
LOCUS 601471858F1 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:3874886 5',
DEFINITION mRNA sequence.
ACCESSION BE783372
VERSION BE783372.1 GI:10204570
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 867)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Place: LLN9633 row: c column: 15
High quality sequence stop: 653.

FEATURES
Location/Qualifiers
1..867
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3874886"
/clone_lib="NIH_MGC_67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
BASE COUNT 253 a 158 c 246 g 210 t
ORIGIN

Query Match 79.5%; Score 620; DB 140; Length 867;
Best Local Similarity 98.6%; Pred. No. 3.2e-175;
Matches 657; Conservative 0; Mismatches 5; Indels 4; Gaps 3;

QY 112 tctaacctaatgctgcaacctggcctcacagcggaagctggaagctggaagagagt 171
|||||
Db 1 TCTAACCTTAATGTCTGCAACCTGGCCTACAGCGGAAGCTGGAAGAGTTGAAGAGAGT 60
|||||
QY 172 attctggcgcgataaatccctggctactgaaactgacagcagcagcagcagcagcagc 231
|||||
Db 61 ATTCTGGCCGATAAATCCCTGGCTACTAGAACTGACCAGGACAGCAGAACTGCATTGCAC 120
|||||
QY 232 tgggcagctgcagctggacatacagaaaattgtggaattttgttgcgaacttggagtgcca 291
|||||
Db 121 TGGGCATGCTCAGCTGGACATACAGAAATGTTGAAATTTTGTGCAACTTGGAGTGCCA 180
|||||
QY 292 gtgaatgataaagacagatgcaggttggctcctcttcataattgcgcttctctgctggccgg 351
|||||
Db 181 GTGAATGATAAGACCATGCGAGTTGGTCTCCTCTTCATATTGCGGCTCTGCTGGCCGG 240
|||||
QY 352 gatgagattgttaaaagcccttctgggaaaaggtgctcaagtgaatgctgtcaatcaaaa 411
|||||
Db 241 GATGAGATTGTAAAGCCCTTCTGGGAAAGGTGCTCAAGTGAATGCTGTCAATCAAAAT 300
|||||
QY 412 ggcgtactccctacattatgcagcttcgaaaaaacagcagcagcagcagcagcagcagc 471
|||||
Db 301 GGCTGTACTCCCTTACATTATGAGCTTCGAAAAACAGGCATGAGATCGTGTCTATGTTA 360
|||||
QY 472 ctggaagcgggggtaataccagatcgaagcattatgagcctacagcagcagcagcagc 531
|||||
Db 361 CTGGAAGCGGGGCTAATCCAGATGCTAAGGACCATATAGGCTACAGCAATGCAACCGG 420
|||||
QY 532 gcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 591
|||||
Db 421 GCAGCAGCCCAAGGTTAACTTGAAGATGATTCAATATCCTTCTGTACTACAAAGCATCCA 480
|||||
QY 592 aacatccaagacactgagggtaaacactcctctacacttagcctgtgatgagagagtg 651
|||||
Db 481 AACATCCAAGACACTGAGGTAACACTCCTCTACACTTACCTTGTGATGAGGAGAGAGTG 540
|||||

```

Qy 652 gaaagcaaaactgctggtgtcccaaggagcaagtatttacattgagaataaagaagaa 711
|||||
Db 541 GAAGACGAAACTG-TGGTGTCCAGGAGCAGATATTACATTGAGATAAGAAGAA 599
|||||
Qy 712 aagacaccctgcaagtggccaaagggtgctggttttaactcaagagaatgggtgaa 771
|||||
Db 600 AAGA-ACCCCTGCAAGTGGC--AAAGGTGGCTGGTGTAACTACTCAAGAGAATGGTGAA 656
|||||
Qy 772 ggttaa 777
|||||
Db 657 GGTAAA 662

RESULT 8
LOCUS BE785737 783 bp mRNA EST 20-OCT-2000
DEFINITION 601478508F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3881328 5',
mRNA sequence.
ACCESSION BE785737
VERSION BE785737.1 GI:10206935
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 783)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9649 row: p column: 01
High quality sequence stop: 665.
FEATURES
source
Location/Qualifiers
1. 783
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3881328"
/clone_lib="NIH_MGC_68"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: Lung; Vector: pCMV-SPOK6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.8 kb. Library constructed by Life
Technologies."
BASE COUNT 218 a 161 c 215 g 189 t
ORIGIN

Query Match 77.7%; Score 605.8; DB 140; Length 783;
Best Local Similarity 96.7%; Pred. No. 5.7e-171;
Matches 672; Conservative 0; Mismatches 17; Indels 6; Gaps 5;

Qy 85 tggacagcgaatggaggggtgtgtctaacctaatggtctgcaacctggcctacagc 144
|||||
Db 1 TGGGACACGGAATGGAGGGGTGTGTCTAACTAAGTGTCTGCAACCTGGCTACAGC 60
|||||
Qy 145 gggaaagctggaagagttgaagagagattctggtccgataaataccctggctactagaact 204
|||||
Db 61 GGGAAAGCTGGGAAGATTGAAGGAGAGATTCTGGCCGATAAATCCCTGGCTACTAGAAGT 120
|||||
Qy 205 gaccagacagcaactgcattgcactggcgcagctgctcagctgacatacagaatgtt 264
|||||
Db 121 GACCAGGACAGCAACTGCATTGCACTGGGCATGCTCAGCTGGACATACAGAAATGTT 180
|||||

```

```

Qy 265 gaattttgttgcacttgagtgccagtgaaatgataaagacgatgcaggttggtctcct 324
|||||
Db 181 GAATTTTGTGTTCAACTTGGAGTGGCAGTGAATGATAAAGACGATGAGTGTGTTCTCT 240
|||||
Qy 325 ctccattatgcgctctctgctgcccggatgagatgtataaagccctcttggaagaggt 384
|||||
Db 241 CTTTCATATTCGGCTTCTGCTGGCCGGATGAGATTGATAAAGCCCTTCGGGAAAAGGT 300
|||||
Qy 385 gctcaagtgaatgctgtcaatcaaaatggcttactcccttaccattatgcagcttcgaaa 444
|||||
Db 301 GCTCAAGTGAATGCTGTCAATCAAAATGGGTGTACTCCCTTACATTATGAGCTTCGAAA 360
|||||
Qy 445 aacaggcatgagatcgctgtcatgttactggaaggcggtgtaataccagatgctaaggac 504
|||||
Db 361 AACAGGCATGAGATCGTGTCTACTGGAAGCGGGCTAATCCAGATGCTAAGGAC 420
|||||
Qy 505 cattatgagctcacagaatgcaccgggagcagcgaaggtaacttgaagatgattcat 564
|||||
Db 421 CATTATGAGGCTACAGCAATGCACCGGGCAGCAGCCAAAGGTAACCTTGAAGATGATTCT 480
|||||
Qy 565 atccttctgtactacaaagcatccacaaacatccaagacactgaggtgaacactcctcta 624
|||||
Db 481 ATCTTCTGTACTACAAAGCATC--ACAAATTTCAAGACACTGAGGGTAACACTCTCTA 538
|||||
Qy 625 cacttagcctgtgatgaggagagtggtggaagaagcaaaactgctggtgtcccaaggagca 684
|||||
Db 539 CACTTAG-CTGTGATGAGGAGAGAGTGGAAAGAGCAAAACTGCTGTTGCC-AGGAGCA 596
|||||
Qy 685 agtattacattgagaataaagaagaagacacccctgcaagtggcgaaggtggcgtg 744
|||||
Db 597 AGTATTTAC-CTGGGAATAAAGAAAGAAAGGACACCTGCAAGGTGGCCAAAGGTGGCC-G 654
|||||
Qy 745 ggttaatactcaagagaatggtggaaggttaaac 779
|||||
Db 655 GGTTAATATCTCCGAGAAATGGTGGACGGTTAAAC 689
|||||

RESULT 9
LOCUS BG568346 634 bp mRNA EST 10-APR-2001
DEFINITION 602587280F1 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:4716163 5',
mRNA sequence.
ACCESSION BG568346
VERSION BG568346.1 GI:13575999
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 634)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.
cDNA Library Preparation: CLONETECH Laboratories, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1563 row: p column: 20
High quality sequence stop: 632.
FEATURES
source
Location/Qualifiers
1. 634
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4716163"
/clone_lib="NIH_MGC_76"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: liver; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccattatggcc); Site_2: SfiI (ggccattatggcc); 5' and

```



```

/clone="2310001K15"
/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
24..719
/notes="putative"
/codon_start=1
/protein_id="BAB26053.1"
/db_xref="GI:12843632"
/translation="MEGCVSNIMICNLAYSGKDELKERILADKSLARTDDSRFAL
HWACSGHFEIVFLLQLQPVNDKDDAGSLHIAASAGDEIVKALVKGHVNAY
NONGCTPLHYAASKRHTAVMLLEGGANPDADHYDATAMHRAAKGNKMKVHILF
YKASNIODEGNTPLHLACDEERVEEAKFLVTQASIVYENKEETPLQVAKGGLG
ILKRLAEGEASM"
polyA_signal      895..900
                    /note="putative"
polyA_site        915
                    /note="putative"
BASE COUNT      249 a 187 c 243 g 236 t
ORIGIN
Query Match      77.4%; Score 603.8; DB 192; Length 915;
Best Local Similarity 91.2%; Pred. No. 2.4e-170;
Matches 641; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
QY 75 aagtagtctggagacagcaaaatgaggggtgtgtgtctctaaactaatggtgcgaacct 134
DB 2  ACCGAGGCTCTGGGAGGCAATGGAGGGGTGTGTCTAACATAATGATCTGTAACTT 61
QY 135 ggcctacagcgggaagctggaagagtggaagagagattctctggccgataaaatccctggc 194
DB 62  GGCTACAGTGGGAGCTGATGATGTGAAGAGCGCATTTTGGCTGATAAATCTCTGGC 121
QY 195 tactagaactgaccagagacagcaagactgcatgcaactggcgatgctcagctggacatac 254
DB 122 TACTAGAACTGATCAGGACAGCAAGACTTTGCACCTGGGCACTGCTCAGCTGGCCATAC 181
QY 255 agaaattgtgaattttgttgaacttgagtcagtcagtgaaatgataaaacacatcagg 314
DB 182 AGAAATTTGTAATCTTGCTGCACTTGAGTGGAGTGCAGTGAATGATAAAGATGACGCAGG 241
QY 315 ttgctctctctctcatattgctgctgctgctgctgctgctgctgctgctgctgcttct 374
DB 242 TTGCTCTCTCTTCATATTGCTGCTCGCTCGCTGCGCGGGATGAGATTGTAAGGCCCTTCT 301
QY 375 gggaaaagggtcctaagtgatgctgctcaatacaaaatggtgttactcccttacattatgc 434
DB 302 GGTGAAAGGTGCACATGTAATGTGTCAATCAAAACGGCTGCACCTCCACTCCATTATGC 361
QY 435 agcttcaaaaacagcatagatcgctgctcatgttactggaagcgggctaatccaga 494
DB 362 ACCTTCGAAGAAATAGGATGATGATGCTGTATGTTACTAGAAAGGTGGGGCTTAACCCAGA 421
QY 495 tgcgaagaccattatgagctacagcaaatgcacggcgagcgagcgcaagggtaacttgaa 554
DB 422 TCGAAGGACCATTTACGATGCTACAGCAATGCACCGGGCAGCAGCCAAAGGTTAACTTGA 481
QY 555 gatgattcatactcttctgactacaagcatcccaaaaacatcccaagacactgagggtaa 614
DB 482 GATGGTTACATCTCTGTCTTCAAAAGGATCCACAAACATCCAAAGACACTGAGGGTAA 541
QY 615 cactctctacacttagctgtgatgagagagtggaagaagcaaaactgctggtgc 674
DB 542 CACTCTCTACACTTAGCTGTGATGAAGAGAGAGTGGAGAGCAAGCAAAATTTCTGTGTAC 601
QY 675 ccaaggagcaagtattacattgagaataagaagaagaagacacccctgcaagtggcga 734
DB 602 TCAAGGAGCAAGTATTTACATTGAGATAAAGAAAGAAAGACACCCCTGCAAGTGGCCAA 661
QY 735 aggtggcctgggttaataactcaagagaatggtggaaggttaa 777
|||||

```

```

Db 662 AGGGGGCCTGGGTTTAATACTCAAGAGACTAGCAGAAGGTGAA 704
RESULT 11
BE904296
LOCUS      601 bp      mRNA      EST      20-OCT-2000
DEFINITION 601494565F2 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3896368 5',
            mRNA sequence.
ACCESSION  BE904296
VERSION     BE904296.1 GI:10396403
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 601)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-remail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Place: LLN9689 row: b column: 17
            High quality sequence stop: 599.
FEATURES             Location/Qualifiers
     source            1..601
     organism="Homo sapiens"
     db_xref="taxon:9606"
     clone="IMAGE:3896368"
     /clone_lib="NIH_MGC_70"
     /tissue_type="epithelioid carcinoma"
     /lab_host="DH10B (phage-resistant)"
     /note="Organ: pancreas; Vector: pCMV-SPORT6; Site:1: NotI;
            Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.
            Average insert size 1.1 kb. Library constructed by Life
            Technologies."
     BASE COUNT      182 a 126 c 155 g 138 t
     ORIGIN
Query Match      77.1%; Score 601; DB 141; Length 601;
Best Local Similarity 100.0%; Pred. No. 1.4e-169;
Matches 601; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 118 ctaatggtctgcaacctggcctacagcgggaagctggaagagtggaaggagattctg 177
DB 1  CTATGGTCTGCAACCTGGCTACAGCGGGAAGCTGGAAGAGTTGAAGGAGAGATTCTG 60
QY 178 gccgataaatccctggctactagaactgaccaggaagcagaactgcatgcaactgggca 237
DB 61  GCGGATAAATCCCTGGCTACTAGAACTGACCGAGCAGACAGAACTGCATTGCACTGGCA 120
QY 238 tgcctagctgacacacagaaattgttgaattttgttgaacttgagtgccagtgaa 297
DB 121 TGTCTAGCTGGACATACAGAAATTTGTGAATTTTGTGCAACTTGGAGTCCAGTGAAT 180
QY 298 gataaagacgatcaggttgctctctctctcatattgctgctgctgctgctgctgctgag 357
DB 181 GATAAAGACGATGCAGGTTGGTCTCTCTCTCATATATGCGGCTTCTGTCGCCGGATGAG 240
QY 358 attgtaaaaacccctcttggaagaagtgctcaagtgaatgctgctcaatcaaatggctgt 417
DB 241 ATTGTAAAAACCCCTCTCTGGGAAAAGGTGCTCAAGTGAATGCTCTCAATCAATAATGGCTGT 300
QY 418 actcccttacatgatcagcttcgaaaaaacagcagcatgagatcgctgctcatgttactgaa 477
DB 301 AC1CCCTTACATATATGCGAGCTTGGAAAAACAGCATGAGATCGCTGCTGCTTACTGAA 360

```


Db	420	CTGGAAGCGGGCTAATCAGATGCTAAGGACCATTTATGAGGCTACAGCAATGCACGG	479
Qy	532	gcagcagcgaaggtaactgaagatgattcatatccttctgtactacaaagcatccaca	591
Db	480	GCAGCAGCCAA-GGTAACCTTGAAGATGATTCATATCTTCTGTACTACAAAGCATCCACA	538
Qy	592	aca-tccagaactgagggtaaacactctctacacttgcctgtgatagga-gagag	649
Db	539	AACATTCGAAGACACTGAGGGTAACACTCTCTACACTTAGCCCTGTGATGAGGACGAGAG	598
Qy	650	tggaagaagcaaaactctggtgtcccaagagcaagattattacatgagaataaagaag	709
Db	599	TGGAGAGCAAGCAAAACTG-TGGTGTCCCAAGAGCAAGT-TTTACATCGAGCAT-CAGAAG	655
Qy	710	aaagacacccctgcaagtgccaaaggtggcctgggttaataactcaagaagaatgg	766
Db	656	AAAAAGACA-CCTTGACGTTGGCAAGGGTGGTCTGGGCTCTCTACCCAGAGAAGGTGG	711
RESULT	15		
ACCESSION	AK018233	1301 bp mRNA	08-FEB-2001
LOCUS	Mus musculus adult male medulla oblongata cDNA, RIKEN full-length	HTC	
DEFINITION	enriched library, clone:6330549L04, full insert sequence.		
ACCESSION	AK018233		
VERSION	AK018233.1	GI:12857851	
KEYWORDS	CAP trapper.		
SOURCE	Mus musculus (strain:C57BL/6J) adult male medulla oblongata cDNA to		
ORGANISM	mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library		
REFERENCE	Eukaryota; Metazoa; Chordata; Cranial; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (sites)		
JOURNAL	Carninci, P. and Hayashizaki, Y.		
REFERENCE	High-efficiency full-length cDNA cloning		
AUTHORS	Methods Enzymol. 303, 19-44 (1999)		
TITLE	2 (sites)		
JOURNAL	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,		
MEDLINE	Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
REFERENCE	Normalization and subtraction of cap-trapper-selected cDNAs to		
AUTHORS	prepare full-length cDNA libraries for rapid discovery of new genes		
TITLE	Genome Res. 10 (10), 1617-1630 (2000)		
JOURNAL	20499374		
MEDLINE	3 (sites)		
REFERENCE	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,		
AUTHORS	Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,		
TITLE	Kikuchi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T.,		
JOURNAL	Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T.,		
MEDLINE	Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,		
REFERENCE	Watahiki, M., Yoneda, Y., Ishikawa, T., Ogawa, K., Tanaka, T.,		
AUTHORS	Matsura, S., Okazaki, Y., Muramatsu, M., Inoue, Y. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format		
JOURNAL	sequencing pipeline with 384 multipillar sequencer		
MEDLINE	Genome Res. 10 (11), 1757-1771 (2000)		
REFERENCE	20530913		
AUTHORS	4 (sites)		
TITLE	The RIKEN Genome Exploration Research Group Phase II Team and		
JOURNAL	FANTOM Consortium.		
MEDLINE	Functional annotation of a full-length mouse cDNA collection		
REFERENCE	Nature 409, 685-690 (2001)		
AUTHORS	5 (bases 1 to 1301)		
TITLE	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,		
JOURNAL	Arakawa, T., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M.,		
MEDLINE	Imotani, K., Hara, A., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F.,		
REFERENCE	Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T.,		
AUTHORS	Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y.,		
TITLE	Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H.,		
JOURNAL	Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,		
MEDLINE	Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,		
REFERENCE	Tanaka, T., Teijima, Y., Toyota, T., Yamamura, T., Yasunishi, A.,		
AUTHORS	Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.		

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2001, 07:46:05 ; Search time 32.29 Seconds
(without alignments)
399.904 Million cell updates/sec

Title: US-09-509-775-2_COPY_14_226

Perfect score: 1093

Sequence: 1 AVSGKLEELKESILADKSLA.....TPLQVAKGGLILKRWEG 213

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_0601.*
1: /SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1093	100.0	226	18	AAW15483 Human P28. Homo s
2	1093	100.0	226	20	AAV02430 Human gankyrin pro
3	1039	95.1	231	20	AAV02432 Rat gankyrin prote
4	1037	94.9	231	20	AAV02431 Mouse gankyrin pro
5	288.5	26.4	1166	22	AAAB7022 Human SPANK. Homo
6	287	26.3	352	21	AAAB11616 D. immitis ankyrin
7	287	26.3	1745	19	AAW70608 Full length ankyri
8	287	26.3	1745	19	AAW76776 D. immitis ankyrin
9	287	26.3	1745	21	AAAB11589 D. immitis ankyrin
10	283.5	25.9	522	22	AAAB6287 Human tankyrase2 c
11	283.5	25.9	1166	22	AAAB66295 Human tankyrase2 T

12	283.5	25.9	1169	22	AAAB66278 Human tankyrase2 r
13	283.5	25.9	1169	22	AAAB66288 Human tankyrase2 c
14	283.5	25.9	1262	22	AAAB66290 Human tankyrase2 c
15	283.5	25.9	1385	22	AAAB66294 Human tankyrase2 T
16	281.5	25.8	673	21	AAAY4403 Human truncated ta
17	281.5	25.8	949	21	AAAY4404 Mouse SPANK. Mus
18	281.5	25.8	991	22	AAAB7023 Human tankyrase I
19	281.5	25.8	1327	21	AAAB7212 Human tankyrase
20	281.5	25.8	1327	21	AAAY4402 Human tankyrase
21	281.5	25.8	1327	22	AAAB66279 Human tankyrase I
22	281	25.7	1181	22	AAAB66297 Drosophila tankyr
23	280.5	25.7	1166	22	AAAY72589 Human tankyrase ho
24	277.5	25.4	1166	21	AAAB7211 Human tankyrase II
25	271	24.8	302	19	AAW70609 Ankyrin protein PB
26	271	24.8	302	19	AAW76777 B. malayi ankyrin
27	271	24.8	302	21	AAAB11590 B. malaya ankyrin
28	269.5	24.7	1030	19	AAW53572 Human myosin light
29	269.5	24.7	1030	19	AAW41378 Human protein p164
30	269	24.6	1074	20	AAAY05734 Human Grb7 effecto
31	268.5	24.6	976	19	AAW53571 Rat p138 protein.
32	268.5	24.6	976	19	AAW41377 Rat protein p138.
33	265	24.2	756	22	AAAB66286 Human tankyrase2 c
34	265	24.2	784	22	AAAB66285 Human tankyrase2 c
35	265	24.2	907	22	AAAB48574 Human breast cance
36	262	24.0	303	19	AAW70606 Ankyrin protein fr
37	262	24.0	303	19	AAW76774 D. immitis ankyrin
38	262	24.0	303	21	AAAB11587 D. immitis ankyrin
39	260.5	23.8	978	21	AAAB42288 Human ORFX ORF7052
40	252	23.1	763	21	AAAY79154 Mouse protein kina
41	252	23.1	786	21	AAAY69163 Amino acid sequenc
42	252	23.1	787	21	AAAY76079 Murine protein kin
43	252	23.1	787	22	AAAB6018 Skin cell protein,
44	250.5	22.9	378	22	AAAB4963 Human secreted pro
45	250	22.9	982	19	AAW71632 Human myosin L-Cha

ALIGNMENTS

RESULT 1	
AAW15483	AAW15483 standard; Protein; 226 AA.
ID	AAW15483
XX	AAW15483;
AC	AAW15483;
XX	17-JUN-1997 (first entry)
DT	Human P28.
XX	Human; proteasome; P28; diagnosis; malignant tumour.
DE	Homo sapiens.
XX	JP09075085-A.
XX	25-MAR-1997.
PD	13-SEP-1995; 95JP-0235052.
XX	13-SEP-1995; 95JP-0235052.
PR	(SAGA) SAGAMI CHEM RES CENTRE.
XX	WPI; 1997-239267/22.
DR	N-PSDB; AAT66424-25.
XX	Human 26S proteasome constituting component protein - useful in the
XX	diagnosis of e.g. malignant tumour
PT	Claim 1; Page 6-7; 9pp; Japanese.
XX	This sequence represents the human proteasome component protein P28.
CC	The protein, P28, is useful for the diagnosis and treatment of
OC	

CC various diseases caused by proteasomes such as malignant tumour.

XX Sequence 226 AA;

Query Match 100.0%; Score 1093; DB 18; Length 226;
Best Local Similarity 100.0%; Pred. No. 1.5e-111;
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYSKLEELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFLQLGVPVNDKDDAG 60
DB 14 AYSKLEELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFLQLGVPVNDKDDAG 73
QY 61 WSPHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHEIAVMLLEGANPD 120
DB 74 WSPHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHEIAVMLLEGANPD 133
QY 121 AKDHYEATAMHRAAKGNLKMTHILLYKASTNIQDTGNTPLHLACDEERVEEAKLLVS 180
DB 134 AKDHYEATAMHRAAKGNLKMTHILLYKASTNIQDTGNTPLHLACDEERVEEAKLLVS 193
QY 181 OGASIYIENKEEKTPLQVAKGGGLILKRMVEG 213
DB 194 OGASIYIENKEEKTPLQVAKGGGLILKRMVEG 226

RESULT 2

AAV02430
ID AAY02430 standard; Protein; 226 AA.
XX
AC AAY02430;
XX
DT 14-JUL-1999 (first entry)
XX
DE Human gankyrin protein.
XX
KW Gankyrin; apoptosis induction; diagnosis; treatment; cancer;
KW hepatocellular carcinoma; oncogenesis mechanism.
XX
OS Homo sapiens.
XX
PN WO9918201-Al.
XX
PD 15-APR-1999.
XX
PF 02-OCT-1998; 98WO-JP04467.
XX
PR 03-OCT-1997; 97JP-0286214.
XX
PA (FUJI/) FUJITA.
XX
PI Fujita J;
XX
DR WPI; 1999-277266/23.
DR N-PSDB; AAX35852.
XX
XX Gankyrin polypeptides, useful for treatment and diagnosis of
PT cancers, e.g. hepatocellular carcinoma, and study of oncogenesis
PT mechanism
XX
PS Claim 1; Page 70-71; 11pp; Japanese.
XX
CC The specification describes human, murine and rat gankyrin DNA and
CC polypeptide sequences. Gankyrin polypeptides inhibit tumorigenic
CC ability and apoptosis induction. The polypeptides and their antibodies
CC can be used in the diagnosis and treatment of cancers,
CC e.g. hepatocellular carcinoma, and study of oncogenesis mechanism.
CC The present sequence represents human gankyrin.
XX
SQ Sequence 226 AA;

Query Match 100.0%; Score 1093; DB 20; Length 226;

Best Local Similarity 100.0%; Pred. No. 1.5e-111;
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYSKLEELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFLQLGVPVNDKDDAG 60
DB 14 AYSKLEELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFLQLGVPVNDKDDAG 73
QY 61 WSPHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHEIAVMLLEGANPD 120
DB 74 WSPHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHEIAVMLLEGANPD 133
QY 121 AKDHYEATAMHRAAKGNLKMTHILLYKASTNIQDTGNTPLHLACDEERVEEAKLLVS 180
DB 134 AKDHYEATAMHRAAKGNLKMTHILLYKASTNIQDTGNTPLHLACDEERVEEAKLLVS 193
QY 181 OGASIYIENKEEKTPLQVAKGGGLILKRMVEG 213
DB 194 OGASIYIENKEEKTPLQVAKGGGLILKRMVEG 226

RESULT 3

AAV02432
ID AAY02432 standard; Protein; 231 AA.
XX
AC AAY02432;
XX
DT 14-JUL-1999 (first entry)
XX
DE Rat gankyrin protein.
XX
KW Gankyrin; apoptosis induction; diagnosis; treatment; cancer;
KW hepatocellular carcinoma; oncogenesis mechanism.
XX
OS Rattus sp.
XX
PN WO9918201-Al.
XX
PD 15-APR-1999.
XX
PF 02-OCT-1998; 98WO-JP04467.
XX
PR 03-OCT-1997; 97JP-0286214.
XX
PA (FUJI/) FUJITA.
XX
PI Fujita J;
XX
DR WPI; 1999-277266/23.
DR N-PSDB; AAX35854.
XX
XX Gankyrin polypeptides, useful for treatment and diagnosis of
PT cancers, e.g. hepatocellular carcinoma, and study of oncogenesis
PT mechanism
XX
PS Claim 1; Page 76-78; 11pp; Japanese.
XX
CC The specification describes human, murine and rat gankyrin DNA and
CC polypeptide sequences. Gankyrin polypeptides inhibit tumorigenic
CC ability and apoptosis induction. The polypeptides and their antibodies
CC can be used in the diagnosis and treatment of cancers,
CC e.g. hepatocellular carcinoma, and study of oncogenesis mechanism.
CC The present sequence represents rat gankyrin.
XX
SQ Sequence 231 AA;

Query Match 95.1%; Score 1039; DB 20; Length 231;

Best Local Similarity 94.8%; Pred. No. 1.2e-105;
Matches 201; Conservative 9; Mismatches 2; Indels 0; Gaps 0;

QY 1 AYSKLEELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFLQLGVPVNDKDDAG 60
DB 14 AYSKLEELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFLQLGVPVNDKDDAG 73

```
QY 61 WSPHLIASAGDEIVKALLGKGAOVNAVNGCTPLHYAASKNRHETAVMLLEGANPD 120
Db 74 WSPHLIASAGDEIVKALLGKGAOVNAVNGCTPLHYAASKNRHETAVMLLEGANPD 133
QY 121 AKDHYEATAMHRAAAKGNLKMTHILLYKASTNIQDTGNTPLHLACDEERVEEAKLLVS 180
Db 134 AKHYDATAMHRAAAKGNLKMTHILLYKASTNIQDTGNTPLHLACDEERVEEAKLLVT 193
QY 181 QCASIYIENKEEKTPLQVAKGGLGILKRMVE 212
Db 194 QGASIYIENKEEKTPLQVAKGGLGILKRIVE 225

RESULT 4
AAY02431
ID AAY02431 standard; Protein; 231 AA.
XX
XX
AC AAY02431;
XX
DT 14-JUL-1999 (first entry)
XX
DE Mouse gankyrin protein.
XX
KW Gankyrin; apoptosis induction; diagnosis; treatment; cancer;
KW hepatocellular carcinoma; oncogenesis mechanism.
XX
OS Mus sp.
XX
PN WO9918201-A1.
XX
PD 15-APR-1999.
XX
PF 02-OCT-1998; 98WO-JP04467.
XX
PR 03-OCT-1997; 97JP-0286214.
XX
PA (FUJI)/FUJITA.
XX
PI Fujita J;
XX
DR WPI; 1999-277266/23.
DR N-PSDB; AAX35853.
XX
XX
Gankyrin polypeptides, useful for treatment and diagnosis of
PT cancers, e.g. hepatocellular carcinoma, and study of oncogenesis
PT mechanism
XX
PS Claim 6; Page 73-74; 11pp; Japanese.
XX
XX The specification describes human, murine and rat gankyrin DNA and
XX polypeptide sequences. Gankyrin polypeptides inhibit tumorigenic
XX ability and apoptosis induction. The polypeptides and their antibodies
XX can be used in the diagnosis and treatment of cancers,
XX e.g. hepatocellular carcinoma, and study of oncogenesis mechanism.
XX The present sequence represents mouse gankyrin.
XX
SQ Sequence 231 AA;

Query Match 94.9%; Score 1037; DB 20; Length 231;
Best Local Similarity 93.9%; Pred. No. 2,1e-105;
Matches 199; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 AYSGLKEELKESILADKSLATRTDQSRALHWACSGAGTEIVEFLQLGVNVNDKDDAG 60
Db 14 aysgkldeiladkslatrtddsrstlhwacsaghteiveflqlgvnvndkddag 73
QY 61 WSPHLIASAGDEIVKALLGKGAOVNAVNGCTPLHYAASKNRHETAVMLLEGANPD 120
Db 74 WSPHLIASAGDEIVKALLGKGAOVNAVNGCTPLHYAASKNRHETAVMLLEGANPD 133
QY 121 AKDHYEATAMHRAAAKGNLKMTHILLYKASTNIQDTGNTPLHLACDEERVEEAKLLVS 180
```

```
Db 134 akhydatamhraaakgnlkmthillykastniqdtgntplhlacdeerveeakflvt 193
QY 181 QCASIYIENKEEKTPLQVAKGGLGILKRMVE 212
Db 194 qgasiiyenkeektplqvakggililkrlae 225

RESULT 5
AAB47022
ID AAB47022 standard; Protein; 1166 AA.
XX
XX AAB47022;
XX
DT 29-MAR-2001 (first entry)
XX
DE Human SPANK.
XX
KW SPANK; SAM; sterile alpha motif; PARP; insulin resistance;
KW poly adenosine diphosphate-ribose polymerase; catalytic domain;
KW ANK; ankyrin repeat; cytosol; insulin-responsive aminopeptidase;
KW IRAP; GLUT4; adipocyte; insulin signalling pathway; hyperlipidaemia;
KW glucose intolerance; atherosclerosis; stroke;
KW obesity; cardiac insufficiency; coronary insufficiency; stroke;
KW high blood pressure; non-insulin dependent diabetes; hypertension;
KW hyperuricemia; Syndrome X; muscular dystrophy; muscle atrophy.
XX
XX Homo sapiens.
XX
OS
XX
FH Key
XX Misc-difference 173..209
XX /note= "Encoded by nucleotides 768..878"
XX
XX Misc-difference 327..362
XX /note= "Encoded by nucleotides 1230..1337"
XX
XX Misc-difference 375..398
XX /note= "Encoded by nucleotides 1374..1445"
XX
XX Misc-difference 482..524
XX /note= "Encoded by nucleotides 1695..1823"
XX
XX Misc-difference 641..677
XX /note= "Encoded by nucleotides 2172..2282"
XX
XX WO200077225-A1.
XX
XX 21-DEC-2000.
XX
XX 09-JUN-2000; 2000WO-US15926.
XX
XX 11-JUN-1999; 99US-0138957.
XX
XX (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX (GEO ) GEN HOSPITAL CORP.
XX
XX Chi N, Lodish HF;
XX
XX WPI; 2001-091404/10.
XX N-PSDB; AAC85294.
XX
XX New insulin signalling protein SPANK, useful for reducing body mass,
XX glucose intolerance or insulin resistance and for preventing or
XX treating obesity-related and muscle-related diseases
XX
XX Claim 3; Fig 3; 65pp; English.
XX
XX This sequence represents human SPANK. The SPANK protein comprises
XX 3 domains:
XX (a) a SAM (sterile alpha motif) domain;
XX (b) a PARP (poly adenosine diphosphate-ribose polymerase) catalytic
XX domain; and
XX (c) an ANK domain composed of ankyrin repeats.
XX SPANK is a cytosolic protein which can poly(ADP-ribosyl)ate itself.
XX SPANK binds insulin-responsive aminopeptidase (IRAP) and modulates
XX translocation of GLUT4 in the perinuclear region of adipocytes. It
XX is an effector in the insulin signalling pathway in eukaryotic cells.
```


(HESK-) HESKA CORP.

Blehm ES, Tang L;

WPI; 2000-375493/32.

N-PSDB; AAA58193, AAA58194, AAA58195, AAA58196.

New Dirofilaria and Brugia ankyrin proteins and nucleic acid encoding them, useful for treating and protecting animals from diseases caused by parasitic helminths, e.g. heartworm disease, elephantiasis or hydrocele

Example 1; Column 117-128; 120pp; English.

The invention relates to ankyrin proteins and nucleic acids from the parasitic helminths Dirofilaria immitis and Brugia malayi. It also relates to antibodies raised against such ankyrin proteins and to compounds that inhibit Dirofilaria or Brugia ankyrin function. Dirofilaria ankyrin cDNAs were isolated from a D. immitis 48 hour L3 cDNA library using PCR primers based on the sequence of the E1 ankyrin from Onchocerca volvulus and the Caenorhabditis elegans ankyrin UNC-44 genes. Brugia ankyrin cDNAs were isolated from a B. malayi adult female cDNA library using D. immitis ankyrin and C. elegans UNC-44 PCR primers. Dirofilaria or Brugia ankyrin proteins and nucleic acids represent novel targets for anti-helminthic vaccines and drugs. Ankyrin nucleic acid molecules, proteins, vaccines and compositions are useful for protecting animals, particularly dogs, from diseases caused by parasitic helminths (e.g., heartworm disease, elephantiasis or hydrocele), as well as for treating the infection. The ankyrin nucleic acid molecules, proteins, vaccines and compositions of the invention are especially useful in treating and preventing infections caused by filarial nematodes (e.g., D. immitis and B. malayi), and ascarid, capillaria, strongylid, strongyloides, trichostrongyle, or trichurid nematodes and are also useful against cestodes and trematodes. The therapeutic compositions may be administered to mammals, including dogs, cats, humans, ferrets, horses, cattle, sheep, and other pets; economic food animals; or zoo animals. The ankyrin nucleic acid molecules, proteins and compounds may also be used as diagnostic reagents to detect infection by parasitic helminths. Prior art anti-helminthic drugs require repeated administration, which often leads to the development of resistant helminth strains that no longer respond to treatment. Such drugs can also cause harmful side effects in the individual being treated, and a number of these drugs can only treat the symptoms of a parasitic disease, being unable to prevent infection by the parasitic helminth. Elucidation of D. immitis and B. malayi ankyrin protein and DNA sequences facilitates the development of agents which inhibit ankyrin-mediated parasite developmental and migratory pathways. Sequence AAB11589 represents full-length D. immitis ankyrin, and sequences AAB11582-B11588, AAB11591-B11614 and AAB11616-B11629 represent D. immitis ankyrin fragments.

Sequence 1745 AA;

Query Match 26.3%; Score 287; DB 21; Length 1745;

Best Local Similarity 36.5%; Pred. No. 5.9e-22;

Matches 73; Conservative 33; Mismatches 84; Indels 10; Gaps 2;

QY 7 EELKESILA-----DKSLATRTDQDSRTALHWACSAGHTEIVEFLLQLGVPNVNDKDDA 59

Db 521 kegqeevaalmdhgtoktlitk---kgftplhaakynlpvaksllertgtpvdiiegkn 577

QY 60 GWSPLHTAASAGRDEIVKALGKGAOVNAVNGQCTPLHYAASKNRHEIAVMLLEGGANP 119

Db 578 qvtplhaahyundkvalllengasahaaakngytpplhaaaknqmdiastillhykana 637

QY 120 DAKDHYEATAMHRAAGNLMKMHILLYKASTNIQTEGNTPLHLACDERVEEAKLLV 179

Db 638 naeskagftplhlaaqeghremaallliengakvgaqargngitpnhlcaqedrsvaeelv 697

QY 180 SOGASITYENKEETPLQVA 199

Db 698 kenaaidpktkagytpplhva 717

RESULT 10

AAB66287 AAB66287 standard; Protein; 522 AA.

XX AC AAB66287;

XX DT 05-APR-2001 (first entry)

XX DE Human tankyrase2 clone protein sequence SEQ ID NO: 99.

XX KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;

XX KW inflammatory disorder.

XX OS Homo sapiens.

XX PN WO200100849-A1.

XX PD 04-JAN-2001.

XX PF 28-JUN-2000; 2000WO-US17827.

XX PR 29-JUN-1999; 99US-0141582.

XX PA (ICOS-) ICOS CORP.

XX PI Christenson E, Demaggio AJ, Goldman PS, McElligott DL;

XX DR WPI; 2001-102896/11.

XX DR N-PSDB; AAF63925.

XX PT New tankyrase2 polypeptides, useful for treating conditions mediated by poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers, inflammatory and autoimmune disorders

XX PS Example 1; Page 156-157; 242pp; English.

XX CC The present invention provides the protein and coding sequence for the human tankyrase2 protein. This is found in two different versions, designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-ribosylation activity and is involved in the modification of TRF1, which is a telomere-specific binding protein. The regulation of telomere length, in which TRF1 has a role, is linked to ageing and cancer. The sequences are useful in the treatment of cancers and inflammatory disorders.

Sequence 522 AA;

Query Match 25.9%; Score 283.5; DB 22; Length 522;

Best Local Similarity 32.1%; Pred. No. 2.4e-22;

Matches 70; Conservative 37; Mismatches 90; Indels 21; Gaps 2;

QY 3 SGKLEELKESILADKSLATRTDQDSRTALHWACSAGHTEIVEFLLQLGVPNVNDKDDAGNS 62

Db 37 ngdvervkrivtpbkvnsrtdagrkstplhfaagfgkrkdvveyllqnganvqardggll 96

QY 63 PLHTAASAGRDEIVKALGKGAOVNAVNGQCTPLHYAASKNRHEIAVMLLEGGANPDAK 122

Db 97 plhnacsfghaevvnlrlrhgadpnrdwnytplheaakgkldvcivllqhgaetir 156

QY 123 D-----HVEATAMHRAAAKGN-LKMHILLYKASTNIQTEGNT 161

Db 157 ntgtraldldpsakavltgkdkdelllesarsgnekmmalltpnvncshadgrkst 216

QY 162 PLHLACDERVEEAKLLVSOGASIYIENKEETPLQVA 199

Db 217 plhlaagynrvkvqlllqhgdvvhakdkgdvlplhna 254

RESULT 11

AAB66295 AAB66295 standard; Protein; 1166 AA.


```
XX AAB66295;
XX DE
XX KW
XX DT 05-APR-2001 (first entry)
XX DE Human tankyrase2 TANK2-SHORT SEQ ID NO: 135.
XX KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
XX KW Inflammatory disorder.
XX OS Homo sapiens.
XX PN WO200100849-A1.
XX PD 04-JAN-2001.
XX PF 28-JUN-2000; 2000WO-US17827.
XX PR 29-JUN-1999; 99US-0141582.
XX PA (ICOS-) ICOS CORP.
XX PI Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX WPI: 2001-102896/11.
XX N-PSDB; AAF63953.
XX New tankyrase2 polypeptides, useful for treating conditions mediated by
XX poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
XX inflammatory and autoimmune disorders -
XX Claim 3; Page 200-203; 242pp; English.
XX The present invention provides the protein and coding sequence for the
XX human tankyrase2 protein. This is found in two different versions,
XX designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
XX polyADP-ribosylation activity and is involved in the modification of
XX TRF1, which is a telomere-specific binding protein. The regulation of
XX telomere length, in which TRF1 has a role, is linked to ageing and
XX cancer. The sequences are useful in the treatment of cancers and
XX inflammatory disorders.
XX Sequence 1166 AA;

Query Match 25.9%; Score 283.5; DB 22; Length 1166;
Best Local Similarity 32.1%; Pred. No. 7.8e-22;
Matches 70; Conservative 37; Mismatches 90; Indels 21; Gaps 2;

QY 3 SKLEELKESILADKSLATRTDODSRTALHWACSAGHTEIVFLLQLGVPVNDKDDAGWS. 62
DB 34 ngdvervrlvtpekvnsrdtagrktphfaagfgrkdvveyllqnganvgarddggll 93

QY 63 PLHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHETAVMLLEGGANPDAK 122
DB 94 plhnacsfghaeavnllrhgadpnardnwnytplheaakgkidvcivllqhgaeptir 153

QY 123 D-----HYEATAMHRAAKGN-LKMIHILLYKASTNIQDTGNT 161
DB 154 ntldgrtalldapsakavltgkdkellesarsgnekmmalltplnvnchsdgrkst 213

QY 162 PLHLACDERVEAKLLVSGASIYIENKEETPLQVA 199
DB 214 plhlaagynrvkivllqhgadvhakdgdvlpplhna 251

RESULT 12
AAB66278
ID AAB66278 standard; Protein; 1169 AA.
XX AC
XX DT 05-APR-2001 (first entry)
XX DE Human tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
XX KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
XX DT 05-APR-2001 (first entry)
```

```
XX Human tankyrase2 related protein sequence SEQ ID NO: 2.
XX DE
XX KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
XX KW Inflammatory disorder.
XX OS Homo sapiens.
XX PN WO200100849-A1.
XX PD 04-JAN-2001.
XX PF 28-JUN-2000; 2000WO-US17827.
XX PR 29-JUN-1999; 99US-0141582.
XX PA (ICOS-) ICOS CORP.
XX PI Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX WPI: 2001-102896/11.
XX N-PSDB; AAF63837.
XX New tankyrase2 polypeptides, useful for treating conditions mediated by
XX poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
XX inflammatory and autoimmune disorders -
XX Disclosure; Page 109-113; 242pp; English.
XX The present invention provides the protein and coding sequence for the
XX human tankyrase2 protein. This is found in two different versions,
XX designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
XX polyADP-ribosylation activity and is involved in the modification of
XX TRF1, which is a telomere-specific binding protein. The regulation of
XX telomere length, in which TRF1 has a role, is linked to ageing and
XX cancer. The sequences are useful in the treatment of cancers and
XX inflammatory disorders.
XX Sequence 1169 AA;

Query Match 25.9%; Score 283.5; DB 22; Length 1169;
Best Local Similarity 32.1%; Pred. No. 7.8e-22;
Matches 70; Conservative 37; Mismatches 90; Indels 21; Gaps 2;

QY 3 SKLEELKESILADKSLATRTDODSRTALHWACSAGHTEIVFLLQLGVPVNDKDDAGWS. 62
DB 37 ngdvervrlvtpekvnsrdtagrktphfaagfgrkdvveyllqnganvgarddggll 96

QY 63 PLHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHETAVMLLEGGANPDAK 122
DB 97 plhnacsfghaeavnllrhgadpnardnwnytplheaakgkidvcivllqhgaeptir 156

QY 123 D-----HYEATAMHRAAKGN-LKMIHILLYKASTNIQDTGNT 161
DB 157 ntldgrtalldapsakavltgkdkellesarsgnekmmalltplnvnchsdgrkst 216

QY 162 PLHLACDERVEAKLLVSGASIYIENKEETPLQVA 199
DB 217 plhlaagynrvkivllqhgadvhakdgdvlpplhna 254

RESULT 13
AAB66288
ID AAB66288 standard; Protein; 1169 AA.
XX AC
XX DT 05-APR-2001 (first entry)
XX DE Human tankyrase2 clone consensus protein SEQ ID NO: 101.
XX KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
```

KW inflammatory disorder.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US17827.
XX
PR 29-JUN-1999; 99US-0141582.
XX
XX (ICOS-) ICOS CORP.
XX
XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX WPI; 2001-102896/11.
XX N-PSDB; AAF63926.
XX
XX New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders -
XX
XX Example 1; Page 162-1665; 242pp; English.
XX
XX The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
CC polyADP-ribosylation activity and is involved in the modification of
CC TRF1, which is a telomere-specific binding protein. The regulation of
CC telomere length, in which TRF1 has a role, is linked to ageing and
CC cancer. The sequences are useful in the treatment of cancers and
CC inflammatory disorders.
XX
XX Sequence 1169 AA;
SQ
Query Match 25.9%; Score 283.5; DB 22; Length 1169;
Best Local Similarity 32.1%; Pred. No. 7.8e-22;
Matches 70; Conservative 37; Mismatches 90; Indels 21; Gaps 2;
QY 3 SGKLEELKESILADKSLATRTDQDSRTALHWACSAGTEIVEFLQLGVVPVNDKDDAGWS 62
Db 37 ngdvervkrlvtpekvnsrdtagrktstplhfaagfgrkdvveyllqnganvqarddggli 96
QY 63 PLHTAASAGRDEIVKALLGKGAQVNAVNONGCTPLHYAASKNRHEIAVMLEGGANPDAK 122
Db 97 plhnacsfghaevvnlllrhgdapnrdwnnytpplheaakgkidvcivilqhgaeptr 156
QY 123 D-----HYEATAMHRAAAKGN-LKMIHILLYKASTNIQDTGNT 161
Db 157 ntgdrtaldldapsakavltgcykddellesarsgnekmmalltpinvnchasdgrkst 216
QY 162 PLHLACDEERVEEAKLLVSGQSIYIENKEETPLQVA 199
Db 217 plhlaagynrvkivqlilqhgdvhdakdglvplhna 254
RESULT 14
AAB66290
ID AAB66290 standard; Protein; 1262 AA.
XX
AC AAB66290;
XX
XX 05-APR-2001 (first entry)
XX
XX Human tankyrase2 clone consensus protein SEQ ID NO: 107.
XX
XX Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder.
XX
XX Homo sapiens.
OS

PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US17827.
XX
PR 29-JUN-1999; 99US-0141582.
XX
XX (ICOS-) ICOS CORP.
XX
XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX WPI; 2001-102896/11.
XX N-PSDB; AAF63930.
XX
XX New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders -
XX
XX Example 2; Page 173-176; 242pp; English.
XX
XX The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
CC polyADP-ribosylation activity and is involved in the modification of
CC TRF1, which is a telomere-specific binding protein. The regulation of
CC telomere length, in which TRF1 has a role, is linked to ageing and
CC cancer. The sequences are useful in the treatment of cancers and
CC inflammatory disorders.
XX
XX Sequence 1262 AA;
SQ
Query Match 25.9%; Score 283.5; DB 22; Length 1262;
Best Local Similarity 32.1%; Pred. No. 8.8e-22;
Matches 70; Conservative 37; Mismatches 90; Indels 21; Gaps 2;
QY 3 SGKLEELKESILADKSLATRTDQDSRTALHWACSAGTEIVEFLQLGVVPVNDKDDAGWS 62
Db 130 ngdvervkrlvtpekvnsrdtagrktstplhfaagfgrkdvveyllqnganvqarddggli 189
QY 63 PLHTAASAGRDEIVKALLGKGAQVNAVNONGCTPLHYAASKNRHEIAVMLEGGANPDAK 122
Db 190 plhnacsfghaevvnlllrhgdapnrdwnnytpplheaakgkidvcivilqhgaeptr 249
QY 123 D-----HYEATAMHRAAAKGN-LKMIHILLYKASTNIQDTGNT 161
Db 250 ntgdrtaldldapsakavltgcykddellesarsgnekmmalltpinvnchasdgrkst 309
QY 162 PLHLACDEERVEEAKLLVSGQSIYIENKEETPLQVA 199
Db 310 plhlaagynrvkivqlilqhgdvhdakdglvplhna 347
RESULT 15
AAB66294
ID AAB66294 standard; Protein; 1385 AA.
XX
AC AAB66294;
XX
XX 05-APR-2001 (first entry)
XX
XX Human tankyrase2 TANK2-LONG SEQ ID NO: 133.
XX
XX Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder.
XX
XX Homo sapiens.
OS
XX WO200100849-A1.
PN
XX PD 04-JAN-2001.
XX

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2001, 07:45:21 ; Search time 19.4 Seconds
(without alignments)
226.069 Million cell updates/sec

Title: US-09-509-775-2_COPY_14_226

Perfect score: 1093

Sequence: 1 AVSGKLEELKESILADKSLA.....TPLOVAKGGLILKRMVEG 213

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA.*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	310	28.4	1088	4	US-09-082-059-2
2	294	26.9	1839	3	US-09-172-977-4
3	287	26.3	352	3	US-09-065-474-139
4	287	26.3	1745	2	US-09-031-485-33
5	287	26.3	1745	2	US-08-847-429A-33
6	287	26.3	1745	3	US-09-065-474-33
7	286	26.2	843	2	US-09-172-977-3
8	271	24.8	302	2	US-09-031-485-38
9	271	24.8	302	3	US-08-847-429A-38
10	271	24.8	302	3	US-09-065-474-38
11	262	24.0	303	2	US-09-031-485-23
12	262	24.0	303	2	US-08-847-429A-23
13	262	24.0	303	3	US-09-065-474-23
14	252	23.1	787	4	US-09-188-930-334
15	242	22.6	1423	4	US-08-810-712-10
16	241	22.0	348	2	US-09-031-485-28
17	241	22.0	348	2	US-08-847-429A-28
18	241	22.0	348	3	US-09-065-474-28
19	237	21.7	387	2	US-08-484-575A-7
20	237	21.7	387	3	US-08-477-459-7
21	237	21.7	387	3	US-08-479-869-7
22	237	21.7	387	4	US-08-486-414-7
23	237	21.7	387	5	PCT-US94-01826A-7
24	237	21.7	387	1	PCT-US94-02252A-7
25	227	20.8	752	1	US-08-281-193-2
26	227	20.8	752	1	US-08-422-106-2
27	227	20.8	752	2	US-08-735-716-2

28	227	20.8	752	2	US-08-555-568B-2	Sequence 2, Appli
29	227	20.8	752	5	PCT-US95-08069-2	Sequence 2, Appli
30	226.5	20.7	131	2	US-09-031-485-20	Sequence 20, Appl
31	226.5	20.7	191	2	US-08-847-429A-20	Sequence 20, Appl
32	226.5	20.7	191	3	US-09-065-474-20	Sequence 2, Appli
33	224.5	20.5	899	1	US-08-365-689-2	Sequence 2, Appli
34	224.5	20.5	899	1	US-08-145-138A-2	Sequence 2, Appli
35	224.5	20.5	933	1	US-07-747-781-2	Sequence 2, Appli
36	224.5	20.5	933	5	PCT-US92-06888-2	Sequence 2, Appli
37	218	19.9	394	2	US-08-555-568B-17	Sequence 17, Appl
38	218	19.9	687	2	US-08-555-568B-21	Sequence 21, Appl
39	218	19.9	688	2	US-08-555-568B-23	Sequence 23, Appl
40	211.5	19.4	452	3	US-09-035-706-2	Sequence 2, Appli
41	211.5	19.4	452	3	US-08-955-841-2	Sequence 2, Appli
42	210	19.2	657	1	US-08-264-534-34	Sequence 34, Appl
43	210	19.2	657	1	US-08-083-590A-13	Sequence 13, Appl
44	210	19.2	657	1	US-08-465-500-34	Sequence 34, Appl
45	210	19.2	657	2	US-08-346-128-34	Sequence 34, Appl

ALIGNMENTS

RESULT 1

US-09-082-059-2
; Sequence 2, Application US/09082059A
; Patent No. 6225086

; GENERAL INFORMATION:

; APPLICANT: Morrow, Jon S.

; APPLICANT: Devorajan, Prasad

; TITLE OF INVENTION: No. 6225086el Ankyrin Proteins and a Method for Their Identifi

; FILE REFERENCE: 44574-5002-US

; CURRENT APPLICATION NUMBER: US/09/082.059A

; CURRENT FILING DATE: 1998-05-21

; EARLIER FILING DATE: 1997-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 1088

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-082-059-2

Query Match 28.4%; Score 310; DB 4; Length 1088;

Best Local Similarity 37.0%; Pred. No. 2.7e-26;

Matches 71; Conservative 36; Mismatches 85; Indels 0; Gaps 0;

Qy 7 EELKESILADKSLATRTDSDSRATLHWACSGAGHTEIVEFLQLGVPVNDKDDAGWSPLHI 66

Db 166 EDVAFLDDHGASLITTKGFTPLRVAAYGKLEVANLLQKSPDACKSGLTPLVH 225

Qy 67 AASAGRDEIVKALGKGAQVNAVNGCTPLHYAASKNRHIEIAVMLEGGANPDADHYE 126

Db 226 AAHYDNQKVALLLDQCGASPHAAKNGYTPHIAAKNQMDIATTLLEYGADANAVTRQG 285

Qy 127 ATAMHRAAAGNKLMIHLLYYKASTNIQDTGNTPLHLACDEERVEAKLLVSGGASII 186

Db 286 IASVHLAAQEGHVDVMSLLGLGRNANVLSNKSGLTPLHLAAQEDRVNVAEYLVNQGAHV 345

Qy 187 IENKEEKTPLQV 198

Db 346 AQTKMGYTPLVH 357

RESULT 2

US-09-172-977-4

; Sequence 4, Application US/09172977

; Patent No. 5989863

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Guegler, Karl J.

APPLICANT: Corley, Neil C.
APPLICANT: Yue, Henry
TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
FILE REFERENCE: PF-0615 US
CURRENT APPLICATION NUMBER: US/09/172,977
CURRENT FILING DATE: 1998-10-14
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PERL Program
SEQ ID NO 4
LENGTH: 1839
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: g29491
US-09-172-977-4

Query Match 26.9%; Score 294; DB 2; Length 1839;
Best Local Similarity 31.9%; Pred. No. 4e-24;
Matches 74; Conservative 40; Mismatches 84; Indels 34; Gaps 2;
QY 1 AVSGKLEELKESILADKSLATRTDSDSTALHWACSAGHTEIVEFLQ----- 48
DB 472 ARAGQVEVVR-CLLRNGALVDARAREEQTPPLHIAISRLGKTEIVQLLQHMHPDAATTNG 530
QY 49 -----LGVVPVNDKDDAGWSPLHIAASAGRDEIVKALLGKGAQVN 87
DB 531 YTPLHISAREGOVDVASVLEAGAAHSLATKGFPLHVAKYGSLDVAKLLQRRRAAD 590
QY 88 AVYNGCCTPLHYAASKNRHEIATVMLEGGANPDADKDHYEATAMHRAAAKGNLKMTHILLY 147
DB 591 SAGKNGLTPLHYAAHYNDGNKALLLEKASPHATAKNGYTPPLHIAAKKNQMDIASTLLN 650
QY 148 YKASTNIQDTGNTPLHLACDEERVEEAKLLVSOGASIVYENKEEKTPLQVA 199
DB 651 YCAETNIVTKQGVTPPLHLSAQEGHDMVTLLLDKGANIHMSTKSLGTSLSHLA 702

RESULT 3
US-09-065-474-139
Sequence 139, Application US/09065474
Patent No. 6063599
GENERAL INFORMATION:
APPLICANT: Tang, Liang
APPLICANT: Blehm, E. Scot
TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
NUMBER OF SEQUENCES: 171
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carol Talkington Verser, Ph.D.
ADDRESSEE: Heska Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,474
FILING DATE: 24-APR-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: HW-5-C1
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505

INFORMATION FOR SEQ ID NO: 139:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-09-065-474-139

Query Match 26.3%; Score 287; DB 3; Length 352;
Best Local Similarity 36.5%; Pred. No. 2e-24;
Matches 73; Conservative 33; Mismatches 84; Indels 10; Gaps 2;
QY 7 BELKESILA-----DKSLATRTDSDSTALHWACSAGHTEIVEFLQLQGVVPVNDKDDA 59
DB 86 KEGQEEVAAAILMDHGTDKTLTK---KGFTPLHIAAKYGNLPAKSLLEGRGTPVDIEGKN 142
QY 60 GWSPLHTAASAGRDEIVKALLGKGAQVNAVQNGCTPLHYAASKNRHEIATVMLEGGANP 119
DB 143 QVTPHLYAAHYNDGNKALLLENGASAAHAAKNGYTPPLHIAAKKNQMDIASTLLHYKANA 202
QY 120 DAKDHYEATAMHRAAAKGNLKMTHILLYKASTNIQDTGNTPLHLACDEERVEEAKLLV 179
DB 203 NAEKAGTTPPLHIAAQEGHREMAALLIENGAKVGAQARNGLTPMLCAQEDRSVAAEELV 262
QY 180 SQGASIVYENKEEKTPLQVA 199
DB 263 KENAAIDPKTRAGYTPHVA 282

RESULT 4
US-09-031-485-33
Sequence 33, Application US/09031485
Patent No. 5824306
GENERAL INFORMATION:
APPLICANT: Tang, Liang
APPLICANT: Blehm, E. Scot
TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carol Talkington Verser, Ph.D.
ADDRESSEE: Heska Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031,485
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION NUMBER: US/08/847,429
FILING DATE: 24-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: HW-5
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 1745 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-09-031-485-33

Query Match 26.3%; Score 287; DB 2; Length 1745;
Best Local Similarity 36.5%; Pred. No. 2.3e-23;
Matches 73; Conservative 33; Mismatches 84; Indels 10; Gaps 2;

QY 7 BELKESILA-----DKSLATRTDQDSRTALHWACSAGHTEIVEFLQLGVVPVNDKDDA 59
DB 521 KEGQEEVAAILMDHGTDKTLTK---KGTPLHLAAKYNLPAKSLLEGGTPVDIEGKN 577
QY 60 GWSPLHIAASAGRDEIVKALLKGQAVNAVNGCTPLHYAASKNRHETAVMLLEGANP 119
DB 578 QVTPLHVAHYNNYNDKVALLENGASAHAAKNGYTPPLHIAAKKNQMDIASTLLHYKANA 637
QY 120 DAKDHYEATAMHRAAKGNLKMHIILLYKASTNTQDTGNTPLHLACDEERVEEAKLLV 179
DB 638 NAEKAGFTPLHLAAQEGHREMAALLIENGAKVGAQARNGLTPMHLCAQEDRVSVAEELV 697
QY 180 SOGASTIYIENKEEKTPLQVA 199
DB 698 KENAAIDPKTKAGYTPPLHVA 717

RESULT 5

US-08-847-429A-33
Sequence 33, Application US/08847429A
Patent No. 5827692

GENERAL INFORMATION:
APPLICANT: Tang, Liang
APPLICANT: Blehm, E. Scot
TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carol Talkington Verser, Ph.D.
ADDRESSEE: Heska Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/847,429A
FILING DATE: 24-APR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: HW-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 1745 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-847-429A-33

Query Match 26.3%; Score 287; DB 2; Length 1745;
Best Local Similarity 36.5%; Pred. No. 2.3e-23;
Matches 73; Conservative 33; Mismatches 84; Indels 10; Gaps 2;

QY 7 BELKESILA-----DKSLATRTDQDSRTALHWACSAGHTEIVEFLQLGVVPVNDKDDA 59
DB 521 KEGQEEVAAILMDHGTDKTLTK---KGTPLHLAAKYNLPAKSLLEGGTPVDIEGKN 577
QY 60 GWSPLHIAASAGRDEIVKALLKGQAVNAVNGCTPLHYAASKNRHETAVMLLEGANP 119
DB 578 QVTPLHVAHYNNYNDKVALLENGASAHAAKNGYTPPLHIAAKKNQMDIASTLLHYKANA 637
QY 120 DAKDHYEATAMHRAAKGNLKMHIILLYKASTNTQDTGNTPLHLACDEERVEEAKLLV 179
DB 638 NAEKAGFTPLHLAAQEGHREMAALLIENGAKVGAQARNGLTPMHLCAQEDRVSVAEELV 697
QY 180 SOGASTIYIENKEEKTPLQVA 199
DB 698 KENAAIDPKTKAGYTPPLHVA 717

RESULT 6

US-09-065-474-33
Sequence 33, Application US/09065474
Patent No. 6063599

GENERAL INFORMATION:
APPLICANT: Tang, Liang
APPLICANT: Blehm, E. Scot
TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 171
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carol Talkington Verser, Ph.D.
ADDRESSEE: Heska Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,474
FILING DATE: 24-APR-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: HW-5-C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 1745 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-065-474-33

Query Match 26.3%; Score 287; DB 3; Length 1745;
Best Local Similarity 36.5%; Pred. No. 2.3e-23;
Matches 73; Conservative 33; Mismatches 84; Indels 10; Gaps 2;

QY 7 BELKESILA-----DKSLATRTDQDSRTALHWACSAGHTEIVEFLQLGVVPVNDKDDA 59
DB 521 KEGQEEVAAILMDHGTDKTLTK---KGTPLHLAAKYNLPAKSLLEGGTPVDIEGKN 577
QY 60 GWSPLHIAASAGRDEIVKALLKGQAVNAVNGCTPLHYAASKNRHETAVMLLEGANP 119
DB 578 QVTPLHVAHYNNYNDKVALLENGASAHAAKNGYTPPLHIAAKKNQMDIASTLLHYKANA 637

TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:

Query Match

```

Query Match      24.08; Score 262; DB 3; Length 303;
Best Local Similarity 33.99; Pred. No. 1.1e-21;
Matches 74; Conservative 35; Mismatches 97; Indels 12; Gaps 4;

QY 1 AYSCKLEBELKESIIADKSLATRTDQDSR-----TALHWACSAGHTEIVEFLQLQGPVND 55
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 16 AHCGHVRVAK--LILDDR-----ADPNARALNGFTPLHIACKNKRIKIVLELLKYHAAIEA 69

QY 56 KDDAGWSPLHIIAASAGRDEIVKALLIGLGAQVNAVQNQGTPLPHYTAASKNRHIEIAMVLEG 115

```


THIS PAGE BLANK (USPTO)

the repeat domain.

A:Reference number: Z22237; MUID:95340633

A:Accession: T42714

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1765 <PET>

A:Cross-references: EMBL:L40632; NID:g710548; PID:g710551; PIDN:AAB01605.1

A:Experimental source: strain C57BL/6J; kidney

C:Genetics:

A:Gene: Ank3

A:Map position: 10

A:Introns: 1587/1

C:Superfamily: ankryrin; ankryrin repeat homology

C:Keywords: alternative splicing

Query Match 28.6%; Score 312.5; DB 2; Length 1765;

Best Local Similarity 35.5%; Pred. No. 4.8e-19;

Matches 72; Conservative 31; Mismatches 67; Indels 33; Gaps 1;

QY 29 TALHWACSGAGTEIVEFLQLQGVPNVDKDDAGWSPLHIAASAGRDEIVKALLGK----- 82

DB 517 TPLHLAAREGHEDVAAPFLDHGASLSITTKGFTPLHVAARYKLEVASLLLOKSASPD 576

QY 83 -----GAOVNAVNGCTPLHYAASKNRHEIAVMLEGG 115

DB 577 AGKSLTPLHVAHYDNQKVALLLDQASPHAAKNGYTPHLHIAAKNQMDIATSLLEY 636

QY 116 GANPDADKHYEATAMHRAAKGNLKMTHILLYKASTNIQDTEGTPHLHACDEERVEEA 175

DB 637 GADANAVTROGIAHVLAHQEGHVDVMSLLSRNANVLSNKSGLTPLHLAAQEDRVNVA 696

QY 176 KLIVSOGASIIYENKEETPLQV 198

DB 697 EVLVNOGAHVDAQTKMGYTPLVH 719

RESULT 3

T42715

ankryrin 3, splice form 3 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000

C:Accession: T42715

R:Peterson, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Vialamas, M.; Turtzo,

J. Cell Biol. 130, 313-330, 1995

A:Title: Ank3 (epithelial ankryrin), a widely distributed new member of the ankryrin gene

the repeat domain.

A:Reference number: Z22237; MUID:95340633

A:Accession: T42715

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1940 <PET>

A:Cross-references: EMBL:L40632; NID:g710548; PID:g710549; PIDN:AAB01604.1

A:Experimental source: strain C57BL/6J; kidney

C:Genetics:

A:Gene: Ank3

A:Map position: 10

A:Introns: 834/1

C:Superfamily: ankryrin; ankryrin repeat homology

C:Keywords: alternative splicing

Query Match

Best Local Similarity 28.6%; Score 312.5; DB 2; Length 1940;

Matches 72; Conservative 31; Mismatches 67; Indels 33; Gaps 1;

QY 29 TALHWACSGAGTEIVEFLQLQGVPNVDKDDAGWSPLHIAASAGRDEIVKALLGK----- 82

DB 517 TPLHLAAREGHEDVAAPFLDHGASLSITTKGFTPLHVAARYKLEVASLLLOKSASPD 576

QY 83 -----GAOVNAVNGCTPLHYAASKNRHEIAVMLEGG 115

DB 577 AGKSLTPLHVAHYDNQKVALLLDQASPHAAKNGYTPHLHIAAKNQMDIATSLLEY 636

QY 116 GANPDADKHYEATAMHRAAKGNLKMTHILLYKASTNIQDTEGTPHLHACDEERVEEA 175

DB 637 GADANAVTROGIAHVLAHQEGHVDVMSLLSRNANVLSNKSGLTPLHLAAQEDRVNVA 696

QY 176 KLIVSOGASIIYENKEETPLQV 198

DB 697 EVLVNOGAHVDAQTKMGYTPLVH 719

RESULT 4

T42713

ankryrin 3, splice form 1 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000

C:Accession: T42713

R:Peterson, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Vialamas, M.; Turtzo,

J. Cell Biol. 130, 313-330, 1995

A:Title: Ank3 (epithelial ankryrin), a widely distributed new member of the ankryrin ge

the repeat domain.

A:Reference number: Z22237; MUID:95340633

A:Accession: T42713

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1943 <PET>

A:Cross-references: EMBL:L40632; NID:g710548; PID:g710550; PIDN:AAB01606.1

A:Experimental source: strain C57BL/6J; kidney

C:Genetics:

A:Gene: Ank3

A:Map position: 10

A:Introns: 855/1

C:Function:

A:Description: supposed to play an important role in the polarized distribution of ma

A:Note: major kidney ankryrin

C:Superfamily: ankryrin; ankryrin repeat homology

C:Keywords: alternative splicing

Query Match

Best Local Similarity 28.6%; Score 312.5; DB 2; Length 1943;

Matches 72; Conservative 31; Mismatches 67; Indels 33; Gaps 1;

QY 29 TALHWACSGAGTEIVEFLQLQGVPNVDKDDAGWSPLHIAASAGRDEIVKALLGK----- 82

DB 517 TPLHLAAREGHEDVAAPFLDHGASLSITTKGFTPLHVAARYKLEVASLLLOKSASPD 576

QY 83 -----GAOVNAVNGCTPLHYAASKNRHEIAVMLEGG 115

DB 577 AGKSLTPLHVAHYDNQKVALLLDQASPHAAKNGYTPHLHIAAKNQMDIATSLLEY 636

QY 116 GANPDADKHYEATAMHRAAKGNLKMTHILLYKASTNIQDTEGTPHLHACDEERVEEA 175

DB 637 GADANAVTROGIAHVLAHQEGHVDVMSLLSRNANVLSNKSGLTPLHLAAQEDRVNVA 696

QY 176 KLIVSOGASIIYENKEETPLQV 198

DB 697 EVLVNOGAHVDAQTKMGYTPLVH 719

RESULT 5

T42716

ankryrin 3, splice form 4 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000

C:Accession: T42716

R:Peterson, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Vialamas, M.; Turtzo,

J. Cell Biol. 130, 313-330, 1995

A:Title: Ank3 (epithelial ankryrin), a widely distributed new member of the ankryrin ge

the repeat domain.

A:Reference number: Z22237; MUID:95340633

A:Accession: T42716

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1961 <PET>
A:Cross-references: EMBL:L40632; NID:g710548; PID:g710552; PIDN:AAB01607.1
A:Experimental source: strain C57BL/6J; kidney
C:Genetics:
A:Gene: Ank3
A:Map position: 10
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing

Query Match 28.6%; Score 312.5; DB 2; Length 1961;
Best Local Similarity 35.5%; Pred. No. 4.8e-19;
Matches 72; Conservative 31; Mismatches 67; Indels 33; Gaps 1;
Qy 29 TALHWACSGHTEIVFEFLQLGVPVNDKDDAGWSPHHTAASAGRDEIVKALLGK----- 82
Db 517 TPLHLAAREGHEDVAFLDHRGASLSITTKGFTPLHVAAYKGVKLEVASLLQLQKASPD 576
Qy 83 -----GAQVNAVNGCTPLHYAASKNRHEIAVMLLG 115
Db 577 AGKSGLTPLHVAHYDNQKVALLLDQGASPHAAKNGYTPHIAAKKNQMDIATSLLEY 636
Qy 116 GANPAKDHYEATAMHRAAKGNLMHILLYKASTNIQDTGEGTTPHLACDEERVEEA 175
Db 637 GADANAVTRQGIASVHLAAQEGHVDVMSLLSRNANVLSNKSGLTPLHLAAQEDRVNVA 696
Qy 176 KLLVSOGASIVTENKEETPLQV 198
Db 697 EVLVNQGAVHDQTKMGYTPPLHV 719

RESULT 6
T50984
related to 26s proteasome subunit p28 [imported] - Neurospora crassa
N:Alternate names: protein B7F18.30
C:Species: Neurospora crassa
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
A:Accession: T50984
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, July 2000
A:Reference number: 225286
A:Status: preliminary
A:Accession: T50984
A:Molecule type: DNA
A:Residues: 1-237 <SCH>
A:Cross-references: EMBL:AL389891; GSPDB:GN00116; NCSP:B7F18.30
A:Experimental source: BAC clone B7F18; strain OR74A
C:Genetics:
A:Gene: NCSP:B7F18.30
A:Map position: 6
A:Introns: 17/1; 25/3; 68/3

Query Match 28.4%; Score 310; DB 2; Length 237;
Best Local Similarity 34.5%; Pred. No. 5.9e-20;
Matches 78; Conservative 30; Mismatches 90; Indels 28; Gaps 4;
Qy 1 AVSGKLEELKESILADKSLATRTDQDSRTALHWACSGHTEIVFEFL-OLGVPVNDKDDA 59
Db 12 ARDGRASIVESLLNANPKLAOKDDGRLPIHWACSYNRKVEVLLVKNQGFDPDVEDDM 71
Qy 60 GWSPLHIAASA-GRDEIVKALLGKAQVNAVNGCTPLHYAASKNRHEIAVMLLG 118
Db 72 GWTTPMISAKVSDAIIDLLSRGADINQTNHQSQTALHTASKNNDIDLAKKLSPMK 131
Qy 119 PD-----AKDHYEATAMHRAAKGNLMHILLYKASTNIQDTGEGTTPHLACDEERVE 173
Db 132 KPAASVVRVYKRGQYPLHRAAAGVSPVNLQLQKASPDAGVTPPLHVAEGHGH 191
Qy 174 EAKLVSOGASI-----YIENKEETPLQV 198
Db 192 AAVALLKAGAEATDKMDGYLALDLAPDKVRRFTEKEAREGIEL 237

RESULT 7
A55575
ankyrin 3, long splice form - human
N:Alternate names: ankyrin G
C:Species: Homo sapiens (man)
C:Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 20-Sep-1999
C:Accession: A55575
R:Kordeli, E.; Lambert, S.; Bennett, V.
J. Biol. Chem. 270, 2352-2359, 1995
A:Title: Ankyrin-G. A new ankyrin gene with neural-specific isoforms localized at the
A:Reference number: A55575; MUID:95138209
A:Accession: A55575
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-4377 <KOR>
A:Cross-references: GB:U13616; NID:g608024; PIDN:AAA64834.1; PID:g608025
C:Genetics:
A:Gene: GDB:ANK3
A:Cross-references: GDB:424503; OMIM:600465
A:Map position: 10q21-10q21
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homol
C:Keywords: alternative splicing; peripheral membrane protein
F:73-105/Domain: ankyrin repeat homology <AN01>
F:106-138/Domain: ankyrin repeat homology <AN02>
F:139-171/Domain: ankyrin repeat homology <AN03>
F:172-200/Domain: ankyrin repeat homology <AN04>
F:201-233/Domain: ankyrin repeat homology <AN05>
F:234-286/Domain: ankyrin repeat homology <AN06>
F:267-299/Domain: ankyrin repeat homology <AN07>
F:300-332/Domain: ankyrin repeat homology <AN08>
F:333-365/Domain: ankyrin repeat homology <AN09>
F:366-398/Domain: ankyrin repeat homology <AN10>
F:399-431/Domain: ankyrin repeat homology <AN11>
F:432-464/Domain: ankyrin repeat homology <AN12>
F:465-497/Domain: ankyrin repeat homology <AN13>
F:498-530/Domain: ankyrin repeat homology <AN14>
F:531-563/Domain: ankyrin repeat homology <AN15>
F:564-596/Domain: ankyrin repeat homology <AN16>
F:597-629/Domain: ankyrin repeat homology <AN17>
F:630-662/Domain: ankyrin repeat homology <AN18>
F:663-695/Domain: ankyrin repeat homology <AN19>
F:696-728/Domain: ankyrin repeat homology <AN20>
F:729-761/Domain: ankyrin repeat homology <AN21>
F:762-794/Domain: ankyrin repeat homology <AN22>
F:795-827/Domain: ankyrin repeat homology <AN23>

Query Match 28.4%; Score 310; DB 2; Length 4377;
Best Local Similarity 37.0%; Pred. No. 2.2e-18;
Matches 71; Conservative 36; Mismatches 85; Indels 0; Gaps 0;
Qy 7 EELKESILADKSLATRTDQDSRTALHWACSGHTEIVFEFLQLGVPVNDKDDAGNSPLHI 66
Db 545 EDVAFLDHRGASLSITTKGFTPLHVAAYKGVKLEVASLLQLQKASPDAGKSGLTPLHV 604
Qy 67 AASGRDEIVKALLGKAQVNAVNGCTPLHYAASKNRHEIAVMLLGEGANPDADKHYE 126
Db 605 AAHYDNQKVALLLDQGASPHAAKNGYTPHIAAKKNQMDIATTLLEYGADANAVTRQG 664
Qy 127 ATAMHRAAKGNLMHILLYKASTNIQDTGEGTTPHLACDEERVEEAKLVSOGASTY 186
Db 665 TASVHLAAQEGHVDVMSLLSRNANVLSNKSGLTPLHLAAQEDRVNVAEVLVNOGAHD 724
Qy 187 IENKEETPLQV 198
Db 725 AQTGMGYTPPLHV 736

RESULT 8
T33631
hypothetical protein F40G9.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans

R.;Tse, W.T.; Menninger, J.C.; Yang-feng, T.L.; Francke, U.; Sahr, K.E.; Lux, S.E.; Wang,
Genomics 10, 858-866, 1991
A:Title: Isolation and chromosomal localization of a novel nonerythrocyte ankyrin gene.
A:Reference number: A40334; MUID:92009921
A:Accession: A40334
A:Molecule type: DNA
A:Residues: 463-474; 'PE','A77-495 <'TS>
A:Cross-references: GB:M37123; NID:gI78647; PIDN:AAA62828.1; PID:gI78648
J:Chan, W.; Kordeli, E.; Bennett, V.
R. Cell Biol. 123, 1463-1473, 1993
A:Title: 440-kD ankyrinB: structure of the major developmentally regulated domain and
A:Reference number: A49462; MUID:94075409
A:Accession: A49462
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-3924 <RES>
A:Cross-references: EMBL:D26634; NID:g406287; PIDN:CAA81387.1; PID:g406288
C:Genetics:
A:Gene: GDB:ANK2
A:Cross-references: GDB:I27607; OMIM:106410
A:Map Position: qg25-qg27
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing
F:2-3924/Product: ankyrin 2, long form #status predicted <MAT>
F:2-1443,3585-3924/Product: ankyrin 2, short form #status predicted <MA2>
F:63-95/Domain: ankyrin repeat homology <AN01>
F:96-128/Domain: ankyrin repeat homology <AN02>
F:129-161/Domain: ankyrin repeat homology <AN03>
F:162-190/Domain: ankyrin repeat homology <AN04>
F:191-223/Domain: ankyrin repeat homology <AN05>
F:232-264/Domain: ankyrin repeat homology <AN06>
F:265-297/Domain: ankyrin repeat homology <AN07>
F:298-330/Domain: ankyrin repeat homology <AN08>
F:331-363/Domain: ankyrin repeat homology <AN09>
F:364-396/Domain: ankyrin repeat homology <AN10>
F:397-429/Domain: ankyrin repeat homology <AN11>
F:430-462/Domain: ankyrin repeat homology <AN12>
F:463-495/Domain: ankyrin repeat homology <AN13>
F:496-528/Domain: ankyrin repeat homology <AN14>
F:529-561/Domain: ankyrin repeat homology <AN15>
F:562-594/Domain: ankyrin repeat homology <AN16>
F:595-627/Domain: ankyrin repeat homology <AN17>
F:628-660/Domain: ankyrin repeat homology <AN18>
F:661-693/Domain: ankyrin repeat homology <AN19>
F:694-726/Domain: ankyrin repeat homology <AN20>
F:727-759/Domain: ankyrin repeat homology <AN21>
F:760-792/Domain: ankyrin repeat homology <AN22>
F:793-825/Domain: ankyrin repeat homology <AN23>

Query Match 26.9%; Score 294; DB 2; Length 3924;
Best Local Similarity 31.9%; Pred. No. 4.8e-17;
Matches 74; Conservative 40; Mismatches 84; Indels 34; Gaps 2;

Qy 1 AVSGKLEELKESILADKSIAVTRTDSDSTALHWACSAHGTEIVFLLQ----- 48

Dd 472 ARAGQVEVR-CLLNGALVDARAEETPLHIASRLGKTETVQLLQHMAHPDAATNG 530

Qy 49 -----LGVPVNDDKGWSPFLHIASAGDEIVKALLGGKAQVN 87

Dd 531 YTPLHSAREGOVDVASVLLEEAGAHSLATRKGTPLPVAARYGYSLDVAKLLLQRRAAD 590

Qy 88 AYVONGCPTLVHAASKNRHEITAVMLLEGGANPDADKHDEVATEAMHRAAAKNLKMTIHILLY 147

Dd 591 SAGKNGLTPLVHAAYHDNQKVALLLEKGA SPATAKNGTYPLHIAKKNMQOIASTLLN 650

Qy 148 XYASTNIQDTGNTPHLHLACDREERYEAKLLYSQGASYIIENKEKTPLOQA 199

Dd 651 YCAETNIVTKCGVTPLHLASQEGTHDMVTLLLDKGANIHMSTKGLSTSUHLA 702

RESULT 10
S57697

hypothetical protein YGR232w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein G8564
C:Species: Saccharomyces cerevisiae
C>Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 23-Mar-2001
C:Accession: S57697; S64556; S63913
R:van der Aart, Q.J.M.; Kleine, K.; Steensma, H.Y.
submitted to the EMBL Data Library, June 1995
A:Description: Sequence analysis of the 43 kb CRM1-YLM9-PET54-SM11-PHO81-YHB4-PFK1 region
A:Reference number: S57680
A:Accession: S57697
A:Molecule type: DNA
A:Residues: 1-228 <VAF>
A:Cross-references: EMBL:X87941; NID:g886908; PIDN:CAA611182.1; PID:g886926
R:van der Aart, Q.J.M.; Steensma, H.Y.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64541
A:Accession: S64556
A:Molecule type: DNA
A:Residues: 1-228 <VAF>
A:Cross-references: EMBL:Z73017; NID:g1323418; PIDN:CAA97260.1; PID:e243668; PID:g132341
R:van der Aart, Q.J.M.; Kleine, K.; Steensma, H.Y.
Yeast 12, 385-390, 1996
A:Title: Sequence analysis of the 43 kb CRM1-YLM9-PET54-DIE2-SM11-PHO81-YHB4-PFK1 region
A:Reference number: S63896; MUID:96267763
A:Accession: S63913
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-228 <VAF>
A:Cross-references: EMBL:X87941; NID:g886908; PIDN:CAA611182.1; PID:g886926
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995
C:Genetics:
A:Map position: 7R
A:Note: YGR232w
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
F:71-103/Domain: ankyrin repeat homology <ANI>

Query Match 26.68; Score 290.5; DB 2; Length 228;
Best Local Similarity 38.6; Pred. No. 2.9e-18;
Matches 76; Conservative 30; Mismatches 80; Indels 11; Gaps 5;
Qy 5 KLEELKESILADKSLATRTDSDRTALHWACSGAGTEIVEFLQLQGVPNV--DKDDAGW 61
Db 17 KVQELLHS---KPSLLLOKQDQGRIPLHWSVSFQAHETSFLLSKMENVNLDIPDDSGW 73
Qy 62 SPLHTAASAGRDEIVKALLGKGAQ--VNAVNGCTPLHYAASKNRHEIAVMLEGGANP 119
Db 74 TPFHTACSVGNLEVVKSLYDRPLKPDKNITNQGVTCLHLAVGKKWFEVSQFLIENGASV 133
Qy 120 DAKDHYEATAMHRAAAGNLMKMIHLL--YYKASTNIQDTGNTPLHLACDEERVEEAKLL 178
Db 134 RIKQKFNQIPLHRAASVGSLSKLIELLGLGKRSVAVNQDQKGTPLFLHALAEHGDAVLL 193
Qy 179 VSQASIV--IENKEK 193
Db 194 VEKYCAEYDLDVNGAK 210

RESULT 11
A57282
ankyrin-related protein unc-44 - Caenorhabditis elegans (fragment)
C:Species: Caenorhabditis elegans
C>Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 20-Sep-1999
C:Accession: A57282
R:Otsuka, A.J.; Franco, R.; Shim, K.H.; Tang, L.Z.; Zhang, Y.Y.; Boontrakulpo
J. Cell Biol. 129, 1081-1092, 1995
A:Title: An ankyrin-related gene (unc-44) is necessary for proper axonal guidance in Cae
A:Reference number: A57282; MUID:95263663
A:Accession: A57282
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-1786 <OTS>
A:Cross-references: GB:U21734; NID:g790607; PIDN:AAA85854.1; PID:g790608
C:Genetics:
A:Gene: unc-44
A:Introns: 97/3; 262/2; 535/1; 749/1; 786/3; 883/3; 1001/2; 1007/2; 1051/3; 1666/1; 1
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homol
F:164-192/Domain: ankyrin repeat homology <AN04>
F:358-390/Domain: ankyrin repeat homology <ANI>
F:391-423/Domain: ankyrin repeat homology <AN11>

Query Match 26.48; Score 288.5; DB 2; Length 1786;
Best Local Similarity 36.08; Pred. No. 5.5e-17;
Matches 76; Conservative 37; Mismatches 89; Indels 9; Gaps 4;
Qy 7 BELKESIL---ADKSLATRTDSDRTALHWACSGAGTEIVEFLQLQGVPNVDKDDAGHSP 63
Db 537 EVAGILLDHNADKTLTK---KGFTPLHLASKYGNLEVRLLERGPVVDIEGKNQVTP 593

Qy 64 LHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHEIAVMLEGGANPDADK 123
Db 594 LHVAHYNNDKAVMLLENGASAKAAKNGYTPHLIAAKKNQMEIATSLLOFKADPNKAS 653

Qy 124 HYEATAMHRAAAGNLMKMIHLLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVSOGA 183
Db 654 RAGFTPLHLSAQEGHKEISGLLIENGSDVGAKANNGLTAMHLCAQEDHVPVQAQILYNGA 713

Qy 184 SIYIENKEEKTPLQVA--KGGIGLILKRWE 212
Db 714 EINSKTNAGYTPPLHVACHFGQLNMV-KFLVE 743

RESULT 12
T15346
elegans ankyrin-related unc-44 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jan-2000
C:Accession: T15346
R:Gattung, S.
submitted to the EMBL Data Library, February 1996
A:Description: The sequence of C. elegans cosmid B0350.
A:Reference number: Z18332
A:Accession: T15346
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1815 <GAT>
A:Cross-references: EMBL:U50071; NID:g1208871; PID:g1208875; PIDN:AAA93445.1; CESP:un
C:Genetics:
A:Gene: CESP:unc-44
A:Introns: 97/3; 262/2; 535/1; 749/1; 786/3; 883/3; 1001/2; 1007/2; 1051/3; 1666/1; 1
C:Superfamily: ankyrin; ankyrin repeat homology
F:358-390/Domain: ankyrin repeat homology <ANR>

Query Match 26.48; Score 288.5; DB 2; Length 1815;
Best Local Similarity 36.08; Pred. No. 5.6e-17;
Matches 76; Conservative 37; Mismatches 89; Indels 9; Gaps 4;
Qy 7 BELKESIL---ADKSLATRTDSDRTALHWACSGAGTEIVEFLQLQGVPNVDKDDAGHSP 63
Db 537 EVAGILLDHNADKTLTK---KGFTPLHLASKYGNLEVRLLERGPVVDIEGKNQVTP 593

Qy 64 LHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHEIAVMLEGGANPDADK 123
Db 594 LHVAHYNNDKAVMLLENGASAKAAKNGYTPHLIAAKKNQMEIATSLLOFKADPNKAS 653

Qy 124 HYEATAMHRAAAGNLMKMIHLLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVSOGA 183
Db 654 RAGFTPLHLSAQEGHKEISGLLIENGSDVGAKANNGLTAMHLCAQEDHVPVQAQILYNGA 713

Qy 184 SIYIENKEEKTPLQVA--KGGIGLILKRWE 212
Db 714 EINSKTNAGYTPPLHVACHFGQLNMV-KFLVE 743

A:Accession: T15347
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-2039 <GAT>
A:Cross-references: EMBL:U50071; NID:g1208871; PID:g1208873; PIDN:AAA93443.1; CESP:unc-44
C:Genetics:
A:Gene: CESP:unc-44
A:Introns: 97/3; 262/2; 535/1; 749/1; 786/3; 883/3; 1001/2; 1045/3; 1660/1; 1719/2; 1971/1
C:Superfamily: ankyrin; ankyrin repeat homology

Query Match	26.4%	Score 288.5;	DB 2;	Length 2039;
Best Local Similarity	36.0%;	pred. No. 6.5e-17;		

Qy 98 HYAASNNRHEIAVMLLEGANPDAKDHYEATAMHRAAAKGNLKMHIHILLYKASTNIQDT 157
| | : : : : : | : : : : : | : : : : : | : : : : :
Db 380 HIACKNHIRVMELLKTKGASIDAVTESGLTPLHVASFNGHLPVKNLLQRCASPNSNV 439
Qy 158 EGNTPHLACDEERVEEAKLLVSQGASIYIENKEEKTPLQVAK--GGIGLILKRMVE 212
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : :
Db 440 KVETPLHMAARAGHTEVAKYLLQNKAKANAKAKDDOTPLRCAARIGHTGMV-KLLIE 495

Search completed: August 13, 2001, 07:46:42
Job time: 139 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2001, 07:44:50 ; Search time 12.84 Seconds
(without alignments)
568.237 Million cell updates/sec

Title: US-09-509-775-2_COPY_14_226

Perfect score: 1093

Sequence: 1 AYSGLLEELKESILADKSLA.....TPLQVAKGGGLILKRMVGE 213

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	294	26.9	3924	1	ANK2_HUMAN
2	290.5	26.6	228	1	YG4X_YEAST
3	287.5	26.3	1862	1	ANK1_MOUSE
4	286.5	26.2	1880	1	ANK1_HUMAN
5	273.5	25.0	234	1	Y057_SCHPO
6	247.5	22.6	1401	1	LATA_LATWA
7	247	22.6	1431	1	DAPK_HUMAN
8	245	22.4	832	1	ANK3_HUMAN
9	237	21.7	752	1	PA26_MOUSE
10	236.5	21.6	439	1	AKR_ARATH
11	232	21.2	751	1	PA26_RAT
12	225	20.6	323	1	ANKH_CHRVI
13	224.5	20.5	898	1	KBF2_HUMAN
14	222.5	20.4	806	1	PA26_HUMAN
15	217.5	19.9	452	1	ILK_MOUSE
16	215	19.7	2703	1	NOTC_DROME
17	214	19.6	2318	1	NTC3_MOUSE
18	212.5	19.4	451	1	ILK_CAVPO
19	211.5	19.4	452	1	ILK1_HUMAN
20	211.5	19.4	452	1	ILK2_HUMAN
21	210	19.2	656	1	FEM1_CAEEL
22	208.5	19.1	1964	1	NTC4_MOUSE
23	207	18.9	2531	1	NTC1_RAT
24	204.5	18.7	2444	1	NTC1_HUMAN
25	202	18.5	679	1	RNSA_MOUSE
26	201	18.4	446	1	BCL3_HUMAN
27	197.5	18.1	2437	1	NOTC_BRARE
28	194	17.7	741	1	RNSA_HUMAN
29	194	17.7	764	1	AKR1_YEAST
30	194	17.7	2524	1	NOTC_XENLA
31	193.5	17.7	984	1	KBF1_CHICK
32	189.5	17.3	347	1	GABC_HUMAN
33	189.5	17.3	383	1	GABB_HUMAN

34	188	17.2	969	1	KBF1_HUMAN	P19838 homo sapien
35	187.5	17.2	500	1	CACT_DROME	Q03017 drosophila
36	187.5	17.2	971	1	KBF1_MOUSE	P25799 mus musculus
37	187	17.1	1429	1	L112_CAEEL	P14585 caenorhabdi
38	186.5	17.1	347	1	GABC_MOUSE	Q00421 mus musculus
39	186.5	17.1	382	1	GABB_MOUSE	Q00420 mus musculus
40	186	17.0	522	1	KBF1_RAT	Q63369 rattus norv
41	183	16.7	168	1	CDN6_HUMAN	P43773 homo sapien
42	182.5	16.7	757	1	HT16_HYDAT	P53356 hydra atten
43	181.5	16.6	642	1	YAZA_SCHPO	Q09701 schizosacch
44	178.5	16.3	414	1	GABD_MOUSE	P81069 mus musculus
45	178.5	16.3	2531	1	NTC1_MOUSE	Q01705 mus musculus

ALIGNMENTS

RESULT 1
ANK2_HUMAN
ID ANK2_HUMAN STANDARD; PRT; 3924 AA.
AC Q01484; Q01485;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ANKYRIN 2 (BRAIN ANKYRIN) (ANKYRIN B) (ANKYRIN, NONERYTHROID).
GN ANK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain stem;
RX MEDLINE=94075409; PubMed=8253844;
RA Chan W., Kordeli E., Bennett V.;
RT "440-kD ankyrinB: structure of the major developmentally regulated domain and selective localization in unmyelinated axons.";
RL J. Cell Biol. 123:1463-1473(1993).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RC TISSUE=Brain stem;
RX MEDLINE=91302466; PubMed=1830053;
RA Otto E., Kunimoto M., McLaughlin T., Bennett V.;
RT "Isolation and characterization of cDNAs encoding human brain ankyrins reveal a family of alternatively spliced genes.";
RL J. Cell Biol. 114:241-253(1991).
RN [3]
RP REVISIONS.
RA Carpenter S.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 463-495 FROM N.A.
RX MEDLINE=92009921; PubMed=1833308;
RA Tse W.T., Menninger J.C., Yang-Feng T.L., Francke U., Sahr K.E., Lux S.E., Ward D.C., Forget B.G.;
RT "Isolation and chromosomal localization of a novel nonerythroid ankyrin gene";
RL Genomics 10:858-866(1991).
CC -!- FUNCTION: ANKYRINS ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTO-SKELETAL ELEMENTS; THEY BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN BAND 4.2, TO NA-K ATPASE, TO THE LYMPHOCTE MEMBRANE PROTEIN GP85, AND TO THE CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN. ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO THE CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN; THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
CC -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: PLASMA MEMBRANE OF NEURONS AS WELL AS GLIAL CELLS THROUGHOUT THE BRAIN.
CC -!- PTM: PHOSPHORYLATED AT MULTIPLE SITES BY DIFFERENT PROTEIN KINASES AND EACH PHOSPHORYLATION EVENT REGULATES THE PROTEIN'S STRUCTURE AND FUNCTION (POTENTIAL).
CC -!- SIMILARITY: CONTAINS 23 ANK REPEATS.

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND VARIANTS.
RC TISSUE=Hematopoietic;
RX MEDLINE=90158830; PubMed=2137557;
RA Lux S.E., John K.M., Bennett V.;
RT "Analysis of cDNA for human erythrocyte ankyrin indicates a repeated
RT structure with homology to tissue-differentiation and cell-cycle
RL control proteins.";
RL Nature 344:36-42(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90175370; PubMed=1699849;
RA Lambert S., Yu H., Prchal J.F., Lawler J., Ruff P., Speicher D.,
RA Cheung M.C., Kan Y.W., Palek J.;
RT "cDNA sequence for human erythrocyte ankyrin.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:1730-1734(1990).
RN [3]
RP VARIANT HS ILE-462
RX MEDLINE=96225450; PubMed=8640229;
RA Eber S.W., Gonzalez J.M., Lux M.L., Scarpa A.L., Tse W.T.,
RA Dornwell M., Herbers J., Kugler W., Oezcan R., Pekrun A.,
RA Gallacher P.G., Schroeter W., Forget B.G., Lux S.E.;
RT "Ankyrin-1 mutations are a major cause of dominant and recessive
RT hereditary spherocytosis.";
RL Nat. Genet. 13:214-218(1996).
CC -1- FUNCTION: ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTOSKELETAL
CC ELEMENTS; BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN BAND 4.2, TO
CC NA-K ATPASE, TO THE LYMPHOCTE MEMBRANE PROTEIN GP85, AND TO THE
CC CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN.
CC ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO THE
CC CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;
CC THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC
CC PLASMA MEMBRANE.
CC -1- ALTERNATIVE PRODUCTS: VARIOUS ISOFORMS OF ANKYRIN ARE PRODUCED BY
CC ALTERNATIVE SPLICING. THE SEQUENCE SHOWN IS THAT OF ANKYRIN
CC VARIANT 2.1.
CC -1- PTM: REGULATED BY PHOSPHORYLATION.
CC -1- PTM: ACYLATED BY PALMITIC ACID GROUP(S).
CC -1- DISEASE: DEFECTS IN ANKI ARE THE CAUSE OF DOMINANT AND RECESSIVE
CC HEREDITARY SPHEROCYTOSIS (HS).
CC -1- SIMILARITY: CONTAINS 23 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: X16609; CAA34610.1; -;
CC DR EMBL: M28880; AAA51732.1; -;
CC DR PIR: S08275; SJHUK.
CC DR PIR: A35049; A35049.
CC DR HSSP: Q00420; LAWG.
CC DR MIM: 182900; -;
CC DR InterPro: IPR000488; -;
CC DR InterPro: IPR000906; -;
CC DR InterPro: IPR002110; -;
CC DR Pfam: PF00791; ZU5; 1.
CC DR Pfam: PF00023; ank; 22.
CC DR Pfam: PF00531; death; 1.
CC DR PROSITE: PS50088; ANK_REPEAT; 20.
CC DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
CC DR PROSITE: PS50017; DEATH_DOMAIN; 1.
CC DR PROSITE: PS50017; Alternative splicing; Repeat; ANK repeat;
CC Cytoskeleton; Lipoprotein; Multigene family; Disease mutation;
CC Phosphorylation; Polymorphism.
CC
CC INIT_MET 0
0

FT	DOMAIN	1	826	89 KDA DOMAIN (ANION EXCHANGE PROTEIN
FT	DOMAIN	827	1381	BINDING DOMAIN).
FT	DOMAIN	1382	1880	62 KDA DOMAIN (SPECTRIN BINDING
FT	DOMAIN			DOMAIN)
FT	REPEAT	43	72	55 KDA REGULATORY DOMAIN (REGULATES
FT	REPEAT	76	105	THE BINDING OF ANKYRIN TO SPECTRIN
FT	REPEAT	109	138	AND THE BAND 3 PROTEIN).
FT	REPEAT	142	171	ANK 1.
FT	REPEAT	173	200	ANK 2.
FT	REPEAT	204	233	ANK 3.
FT	REPEAT	237	266	ANK 4.
FT	REPEAT	270	299	ANK 5.
FT	REPEAT	303	332	ANK 6.
FT	REPEAT	336	365	ANK 7.
FT	REPEAT	369	398	ANK 8.
FT	REPEAT	402	431	ANK 9.
FT	REPEAT	435	464	ANK 10.
FT	REPEAT	468	497	ANK 11.
FT	REPEAT	501	530	ANK 12.
FT	REPEAT	534	563	ANK 13.
FT	REPEAT	567	596	ANK 14.
FT	REPEAT	600	629	ANK 15.
FT	REPEAT	633	662	ANK 16.
FT	REPEAT	666	695	ANK 17.
FT	REPEAT	699	728	ANK 18.
FT	REPEAT	732	761	ANK 19.
FT	REPEAT	765	794	ANK 20.
FT	REPEAT	1402	1486	ANK 21.
FT	REPEAT	1512	1873	ANK 22.
FT	REPEAT	1874	1874	ANK 23.
FT	REPEAT	1849	1880	DEATH DOMAIN.
FT	REPEAT			MISSING (IN ISOFORM 2.2).
FT	REPEAT			H -> D (IN ISOFORM 2.2).
FT	REPEAT			TVEGLEPDESELEVDIDYEMKHSKDHSTSPNP -> ELRGS
FT	REPEAT			GLQPDLEGKGAQIVKRSALRGKQ (IN A THIRD
FT	REPEAT			ISOFORM).
FT	REPEAT			R -> T.
FT	REPEAT			/FTID-VAR_000595.
FT	REPEAT			V -> I (IN HS).
FT	REPEAT			/FTID-VAR_000596.
FT	REPEAT			R -> H (IN BRUEGGEN).
FT	REPEAT			/FTID-VAR_000597.
FT	REPEAT			V -> A.
FT	REPEAT			/FTID-VAR_000598.
FT	REPEAT			D -> E.
FT	REPEAT			/FTID-VAR_000599.
FT	REPEAT			S -> T.
FT	REPEAT			E -> D.
FT	REPEAT			/FTID-VAR_000600.
FT	REPEAT			/FTID-VAR_000601.
FT	REPEAT			D -> N (IN DUESSELDORF).
FT	REPEAT			/FTID-VAR_000602.
FT	REPEAT			R -> D.
FT	REPEAT			/FTID-VAR_000603.
FT	REPEAT			A -> S (IN REF. 2).
FT	REPEAT			V -> I (IN REF. 2).
FT	REPEAT			SEQUENCE 1880 AA; 206145 MW; 1C5F5FEFDICD428 CRC64;
FT	REPEAT			

Query Match 26.2%; Score 286.5; DB 1; Length 1880;
Best Local Similarity 37.5%; Pred. No. 1.4e-17;
Matches 72; Conservative 28; Mismatches 83; Indels 9; Gaps 2;
QY 13 ILADKSLATRTDQDSR-----TALHWACASAGHTEIVEFFLLQLGVVNDKDDAGNSPLHIA 67
Db 355 VLLDKG----AKPNSRALNGFTPLHIAKKNHVRVMEILLKTGASIDAVTESGLTPLHVA 410
QY 68 ASAGRDEIVKALLCKGAQVNAVNGCTPLHYAASKNRHETAYVLLLEGGANPDAAKHVEA 127
Db 411 SFMGHLPIVKNLLQRGASPNVSNVYETPLHMAARAGHTEVAKYLLQNKAKVNAKAKDDQ 470
QY 128 TAMHRAAKGNLKMHIHLLVYKASTNIQDTGNTPLHACDEERVEAKLLVSGASIYI 187
INIT_MET 0

Db 792 KTSNGATPLHLATFKGSOAALLLNNEVNRDTEGOMPIGHAAMTGLDLVAQAIIISI 851
QY 149 KAS-TNIQTEGWTPLHLACDEEVEEAKLLVSOGASIIYIENKEETPLQV--AKGGLGL 205
Db 852 DATVVDIEDKNSDTPLNLAAQNSHIDVIKIFIDQAGADINRNKGLAPLAFSKKGNLDM 911
QY 206 I 206
Db 912 V 912
RESULT 7
DAPK_HUMAN
ID DAPK_HUMAN STANDARD; PRT: 1431 AA.
AC P53355;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE DEATH-ASSOCIATED PROTEIN KINASE 1 (EC 2.7.1.-) (DAP KINASE 1).
GN DAPK1 OR DAPK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RZ SEQUENCE FROM N.A.
RP MEDLINE=95129831; PubMed=7828849;
RA Daiss L.P., Feinstein E., Berissi H., Cohen O., Kimchi A.;
RT "Identification of a novel serine/threonine kinase and a novel 15-kD
RT protein as potential mediators of the gamma interferon-induced cell
RT death.";
RL Genes Dev. 9:15-30(1995).
RN [2]
RP REVISIONS TO 164-171.
RA Feinstein E.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN MEDIATING INTERFERON-GAMMA-INDUCED CELL
CC DEATH.
CC -!- PTM: AUTOPHOSPHORYLATED.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 10 ANK REPEATS.
CC -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X76104; CAA53712.1; -
DR HSSP; Q63450; IAO6.
DR MIN; 600831; -
DR InterPro; IPR000488; -
DR InterPro; IPR000719; -
DR InterPro; IPR002110; -
DR InterPro; IPR002290; -
DR Pfam; PF00023; ank; 8.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00069; pkinase; 1.
DR PROSITE; PS50088; ANK_REPEAT; 6.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW transferase; Serine/threonine-protein kinase; Calmodulin-binding;
KW phosphorylation; ATP-binding; Repeat; ANK repeat; Apoptosis.
FT DOMAIN 13 266 PROTEIN KINASE.
FT DOMAIN 267 334 CALMODULIN-BINDING.
FT REPEAT 378 407 ANK 1.
FT REPEAT 411 440 ANK 2.

FT REPEAT 444 473 ANK 3.
FT REPEAT 478 507 ANK 4.
FT REPEAT 511 540 ANK 5.
FT REPEAT 544 573 ANK 6.
FT REPEAT 577 606 ANK 7.
FT REPEAT 610 639 ANK 8.
FT REPEAT 610 639 ANK 9.
FT REPEAT 876 905 ANK 10.
FT REPEAT 1163 1197 ANK 10.
FT DOMAIN 1313 1397 DEATH.
FT NP_BIND 19 27 ATP (BY SIMILARITY).
FT BINDING 42 42 ATP (BY SIMILARITY).
FT ACT_SITE 139 139 BY SIMILARITY.
FT MUTAGEN 42 42 K->A: LOSS OF ACTIVITY.
SQ SEQUENCE 1431 AA; 160017 MW; 9EB4811004A155B CRC64;
Query Match 22.6%; Score 247; DB 1; Length 1431;
Best Local Similarity 29.0%; Pred. No. 3.4e-14;
Matches 61; Conservative 33; Mismatches 82; Indels 34; Gaps 1;
QY 24 DQDSRTALHWACSGAGTEIVEFLQLGVVNDKDDAGWSPLHIAASAGRDE----- 74
Db 409 DRGGSNAVYWAARHGHDVTLKELSENKCPDLVKKSGEMALHVAARYGHADVAQVTCRAS 468
QY 75 -----IVKALLGKGAQYNAVNGCTPLHYAASKNRHETA 109
Db 469 AQIPISRTKEETPLHCAAWHGYYSVAKALCBAGCNVNIKNREGETPLLTASARGYHDIV 528
QY 110 VMLLEGANPDADKHYEATAMHRAAAKGLMKHILLYKASTNIQDTEGNTPLHLACDE 169
Db 529 ECLAHGADLNACDKDGHIALHLAVRCQMEVIKTLLSQGCFVDYQDRHGNTPLHVACKD 588
QY 170 ERVEEAKLLVSOGASIIYIENKEETPLQVA 199
Db 589 GNPPIVVALCEANCLDISNKYGRTPHLA 618
RESULT 8
ANR3_HUMAN
ID ANR3_HUMAN STANDARD; PRT: 832 AA.
AC P57078;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE SERINE/THREONINE-PROTEIN KINASE ANKR3 (EC 2.7.1.-) (ANKYRIN REPEAT
DE DOMAIN PROTEIN 3).
GN ANKR3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RZ SEQUENCE FROM N.A.
RP MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shlimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Lehrach H., Reinhardt R., Vaspo M.-L.;
RT "The DNA sequence of human chromosome 21.";
RL Nature 405:311-319(2000).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 10 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----


```
RC STRAIN=D / ATCC 17899 / DSM 180;
RX MEDLINE=93300842; PubMed=8390993;
RA Dolata M.M., van Beeumen J.J., Ambler R.P., Meyer T.E.,
RA Cusanovich M.A.;
RT "Nucleotide sequence of the heme subunit of flavocytochrome c from
RT the purple phototrophic bacterium, Chromatium vinosum. A 2.6-kilobase
RT pair DNA fragment contains two multitheme cytochromes, a flavoprotein,
RT and a homolog of human ankryrin."
RL J. Biol. Chem. 268:14426-14431(1993).
CC -|- SIMILARITY: CONTAINS 8 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L13419; AAA23315.1; -.
CC HSSP; Q00420; IAWC.
CC InterPro; IPR002110; -.
CC Pfam; PF00023; ank; 8.
CC PROSITE; PS50088; ANK_REPEAT; 7.
CC PROSITE; PS50297; ANK_REPEAT_REGION; 1.
CC Signal; ANK repeat; Repeat.
CC SIGNAL; 1 28 POTENTIAL.
CC CHAIN 29 323 ANKYRIN HOMOLOG.
CC REPEAT 31 60 ANK 1.
CC REPEAT 64 93 ANK 2.
CC REPEAT 97 126 ANK 3.
CC REPEAT 130 159 ANK 4.
CC REPEAT 163 192 ANK 5.
CC REPEAT 196 225 ANK 6.
CC REPEAT 229 258 ANK 7.
CC REPEAT 262 291 ANK 8.
CC SEQUENCE 323 AA; 33453 MW; 3976D34A9BD607DF CRC64;
CC
CC Query Match 20.6%; Score 225; DB 1; Length 323;
CC Best Local Similarity 29.7%; Pred. No. 5.3e-13;
CC Matches 65; Conservative 34; Mismatches 82; Indels 38; Gaps 2;
CC
CC QY 19 LATRTDQSRTA-----LHWACSGHTEIVEFLQLGVVNDKDGWSPLHIAASGRD 73
CC I: |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::||
CC D8 85 LSRGADVARTVAGCTPLTFAEAGHIGISALLLGERGARVHDTRSGWDMIAERHGIT 144
CC QY 74 EIVKALLKGAGVNAVNGCTPLHYAASKNRHETAVMLLEGANPDADKHVEATAMHRA 133
CC ::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::||
CC D8 145 DMVEQLFKGADPKAADREGRTALMQAASKGTGVLPLLIEGGADLEARDKQATALLIA 204
CC QY 134 A-----AKGNLKMHIILLYKASTNIQDTSGN 160
CC |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::||
CC D8 205 ADQOGAGAVETLAGIQAQDAVDALGSTALILVAGHGVAMVERLLAMGADPNRQDRGT 264
CC QY 161 TPLHLACDEERVEEAKLLVSQAGSIYENKEKTPLOVA 199
CC |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::||
CC D8 265 TALMEAVATDHAELDRLIAGARTDLKDDADRTAADIA 303
CC
CC RESULT 13
CC ID KBF2_HUMAN
CC AC Q00653; STANDARD; PRT; 898 AA.
CC DT 01-OCT-1993 (Rel. 27, Created)
CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
CC DT 01-OCT-2000 (Rel. 40, Last annotation update)
CC DE NUCLEAR FACTOR NF-KAPPA-B p100 SUBUNIT (H2TF1) (ONCOGENE LYT-10)
CC DE (LYT10) [CONTAINS: NUCLEAR FACTOR NF-KAPPA-B P52 SUBUNIT].
CC OS Homo sapiens (Human).
CC GN Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC
```

```
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leukemia;
RX MEDLINE=91343004; PubMed=1876189;
RA Schmid R.M., Perkins N.D., Duckett C.S., Andrews P.C., Nabel G.J.;
RT "Cloning of an NF-kappa B subunit which stimulates HIV transcription
RT in synergy with p65."
RL Nature 352:733-736(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92123193; PubMed=1531086;
RA Bours V., Burd P.R., Brown K., Villalobos J., Park S., Ryseck R.P.,
RA Bravo R., Kelly K., Siebenlist U.;
RT "A novel mitogen-inducible gene product related to p50/p105-NF-kappa
RT B participates in transactivation through a kappa B site."
RL Mol. Cell. Biol. 12:685-695(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=94309910; PubMed=8036016;
RA Thakur S., Lin H.C., Tseng W.T., Kumar S., Bravo R., Foss F.,
RA Gelinas C., Rabson A.B.;
RT "Rearrangement and altered expression of the NFKB-2 gene in human
RT cutaneous T-lymphoma cells."
RL Oncogene 9:2335-2344(1994).
RN [4]
RP SEQUENCE OF 1-220 FROM N.A.
RX MEDLINE=95059001; PubMed=7969113;
RA Liptay S., Schmid R.M., Nabel E.G., Nabel G.J.;
RT "Transcriptional regulation of NF-kappa B2: evidence for kappa B-
RT mediated positive and negative autoregulation."
RL Mol. Cell. Biol. 14:7695-7703(1994).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 37-327.
RX MEDLINE=98046037; PubMed=9384586;
RA Cramer P., Larson C.J., Verdine G.L., Mueller C.W.;
RT "Structure of the human NF-kappaB p52 homodimer-DNA complex at 2.1-A
RT resolution."
RL EMBO J. 16:7078-7090(1997).
CC -|- FUNCTION: P100 IS THE PRECURSOR OF THE P52 SUBUNIT OF THE NUCLEAR
CC FACTOR NF-KAPPA-B, WHICH BINDS TO THE KAPPA-B CONSENSUS SEQUENCE
CC 5'-GGGNNYYCC-3', LOCATED IN THE ENHANCER REGION OF GENES INVOLVED
CC IN IMMUNE RESPONSE AND ACUTE PHASE REACTIONS. THE PRECURSOR
CC PROTEIN ITSELF DOES NOT BIND TO DNA.
CC -|- SUBUNIT: ACTIVE NF-KAPPA-B IS A HETERODIMER OF AN ABOUT 52 KDA
CC DNA-BINDING SUBUNIT AND THE WEAK DNA-BINDING SUBUNIT P65. TWO
CC HETERODIMERS MIGHT FORM A LABILE TETRAMER.
CC -|- SUBCELLULAR LOCATION: NUCLEAR, BUT ALSO FOUND IN THE CYTOPLASM
CC IN AN INACTIVE FORM COMPLEXED TO AN INHIBITOR (I-KAPPA-B).
CC -|- ALTERNATIVE PRODUCTS: NFKB2 CAN PRODUCE BOTH A 100 KDA AND A 49
CC KDA PRODUCT FROM AN ALTERNATIVELY SPLICED TRANSCRIPT.
CC -|- DOMAIN: THE C-TERMINUS OF P100 MIGHT BE INVOLVED IN CYTOPLASMIC
CC RETENTION, INHIBITION OF DNA-BINDING BY P52 HOMODIMERS, AND/OR
CC TRANSCRIPTION ACTIVATION (BY SIMILARITY).
CC -|- DISEASE: IN CUTANEOUS T-CELL LEUKEMIA (CTCL) CELL LINE HUT 78, A
CC REARRANGED NFKB2 GENE IS FOUND WHICH ENCODES FOR A TRUNCATED 80
CC KDA PROTEIN (P80HT).
CC -|- SIMILARITY: BELONGS TO THE REL/DORSAL FAMILY.
CC -|- SIMILARITY: CONTAINS 7 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X61498; CAA43715.1; -.
CC D8 EMBL; S76638; AAB21124.1; -.
CC D8 EMBL; U09609; AAA21462.1; -.
CC D8 EMBL; U20816; AAA68171.1; -.
CC D8 PIR; S17233; S17233.
```

PDB; 1A30; 27-MAY-98.
DR TRANSFAC: T01927; -
DR MIM; 164012; -
DR InterPro: IPR000451; -
DR InterPro: IPR000488; -
DR InterPro: IPR002110; -
DR Pfam; PF00354; RHD; 1.
DR Pfam; PF00023; ank; 6.
DR Pfam; PF00531; death; 1.
DR PRINTS; PRO0057; NFKB1NSCPFT.
DR PROSITE; PS0088; ANK_REPEAT; 5.
DR PROSITE; PS0297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS01204; REL_1; 1.
DR PROSITE; PS0254; REL_2; 1.
KW DNA-binding; Transcription regulation; Activator; Nuclear protein;
KW Repeat; ANK repeat; Phosphorylation; Alternative splicing;
KW Disease mutation; Proto-oncogene; 3D-structure.
FT DOMAIN 38 343
FT REPEAT 486 518 ANK 1.
FT REPEAT 525 554 ANK 2.
FT REPEAT 558 590 ANK 3.
FT REPEAT 598 627 ANK 4.
FT REPEAT 632 662 ANK 5.
FT REPEAT 666 695 ANK 6.
FT REPEAT 728 757 ANK 7.
FT DOMAIN 337 341
FT DOMAIN 350 400
FT VARIANT 666
FT VARIANT 667 898
FT CONFLICT 144
FT CONFLICT 213
FT CONFLICT 396
FT CONFLICT 433
FT CONFLICT 455
FT CONFLICT 740
FT CONFLICT 858
FT CONFLICT 874
FT CONFLICT 889
FT CONFLICT 898
SQ SEQUENCE 898 AA; 96752 MW; DFD153E9E23A42B CRC64;
Query Match 20.5%; Score 224.5; DB 1; Length 898;
Best Local Similarity 33.9%; Pred. No. 2e-12;
Matches 65; Conservative 29; Mismatches 85; Indels 13; Gaps 5;
Qy 16 DKSLATRTDQDSRTALHWACSGAHTIEVFLQLQGVFVNDKDDAGWSPLHIA--ASAGR 73
Db 515 DLGVVNLTHLHQTPLHLAVITGTSVSVFLRVGADPALLDRHGDGSAMHLALRAGAGAP 574
Qy 74 BIVKALLGKAQVNAVQ-----NCTPLHYAASKNRHEIAVNLGGANPDADKHYE 126
Db 575 ELLRALLQSGAP--AVPQLLHMPDFEGLYPVHLAVARSPECLDLLVDSCAEVATERQ 632
Qy 127 A-TAMHRAAAGNLKMI-HILLYYKASTNIQDTEGNTPLHLACDEERVEAKLLVSGAS 184
Db 633 GRTALHLATEMEELGLVTHLYTKLRANVARTFAGNTPLHLAAGLYPTLRLLLKAGAD 692
Qy 185 IYIENKEETPL 196
Db 693 IHAENEPLCPL 704
RESULT 14
ID PA26_HUMAN STANDARD; PRT; 806 AA.
AC O60733; O75645; Q9Y671; Q9UIT0; Q9UG29;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)

01-OCT-2000 (Rel. 40, Last annotation update)
85 KDA CALCIUM-INDEPENDENT PHOSPHOLIPASE A2 (EC 3.1.1.4) (IPLA2) (CAI-PLA2).
PLA2G6 OR IPLA2.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A. (ISOFORMS LH-IPLA; ANKYRIN-IPLA2-1 AND -2).
TISSUE-B-cell, and Testis;
MEDLINE=98079046; PubMed=9417066;
Larsson P.K.A., Claesson H.-E., Kennedy B.P.;
"Multiple splice variants of the human calcium-independent phospholipase A2 and their effect on enzyme activity.";
J. Biol. Chem. 273:207-214(1998).
[2]
SEQUENCE FROM N.A. (ISOFORMS LH-IPLA AND SH-IPLA).
TISSUE-Pancreatic islets;
MEDLINE=99194813; PubMed=10092647;
Ma Z., Wang X., Nowatzke W., Ramanadham S., Turk J.;
"Human pancreatic islets express mRNA species encoding two distinct catalytically active isoforms of group VI phospholipase A2 (IPLA2) that arise from an exon-skipping mechanism of alternative splicing of the transcript from the IPLA2 gene on chromosome 22q13.1.";
J. Biol. Chem. 274:9607-9616(1999).
[3]
SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
MEDLINE=99269033; PubMed=10336645;
Larsson Forsell P.K.A., Kennedy B.P., Claesson H.-E.;
"The human calcium-independent phospholipase A2 gene. Multiple enzymes with distinct properties from a single gene.";
Eur. J. Biochem. 262:575-585(1999).
[4]
SEQUENCE FROM N.A. (ISOFORM LH-IPLA).
Graham D.;
Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
[5]
SEQUENCE FROM N.A. (ISOFORM LH-IPLA).
TISSUE-Testis;
Ansorge W., Winkler U., Mewes H.-W., Gassenhuber J., Wiemann S.;
Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
-!- FUNCTION: CATALYZES THE RELEASE OF FATTY ACIDS FROM PHOSPHOLIPIDS. IT HAS BEEN IMPLICATED IN NORMAL PHOSPHOLIPID REMODELLING, NITRIC OXIDE-INDUCED OR VASOPRESSIN-INDUCED ARACHIDONIC ACID RELEASE AND IN LEUKOTRIENE AND PROSTAGLANDIN PRODUCTION. MAY PARTICIPATE IN FAS MEDIATED APOPTOSIS AND IN REGULATING TRANSMEMBRANE ION FLUX IN GLUCOSE-STIMULATED B-CELLS.
-!- FUNCTION: ISOFORMS ANKYRIN-IPLA2-1 AND -2, WHICH LACK THE CATALYTIC DOMAIN, ARE PROBABLY INVOLVED IN THE NEGATIVE REGULATION OF IPLA2 ACTIVITY.
-!- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERYL-PHOSPHOCHOLINE + A FATTY ACID ANION.
-!- SUBUNIT: FORMS LARGE OLIGOMERIC 270-350 KDA STRUCTURES.
-!- SUBCELLULAR LOCATION: THE LONG ISOFORM WAS FOUND TO BE MEMBRANE BOUND. THE SHORT ISOFORM IS CYTOPLASMIC.
-!- ALTERNATIVE PRODUCTS: 4 ISOFORMS; LH-IPLA2 (SHOWN HERE), SH-IPLA2, ANKYRIN-IPLA2-1 AND ANKYRIN-IPLA2-2; ARE PRODUCED BY ALTERNATIVE SPLICING.
-!- TISSUE SPECIFICITY: FOUR DIFFERENT TRANSCRIPTS WERE FOUND TO BE EXPRESSED IN A DISTINCT TISSUE DISTRIBUTION.
-!- SIMILARITY: CONTAINS 7 ANK REPEATS.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF064594; AAC97486.1; -
EMBL; AF102988; AAD41722.1; -

[illegible]

```

RESULT 15
ILK_MOUSE
ID ILK_MOUSE STANDARD; PRT; 452 AA.
AC O55222;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE INTEGRIN-LINKED PROTEIN KINASE [EC 2.7.1.-].
GN ILK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RS SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=Heart;
RX MEDLINE=98031580; PubMed=9366252;
RA Li F., Liu J., Mayne R., Wu C.;
RT "Identification and characterization of a mouse protein kinase that is
RL highly homologous to human integrin-linked kinase.";
CC -! FUNCTION: RECEPTOR-PROXIMAL PROTEIN KINASE REGULATING INTEGRIN-
CC MEDIATED SIGNAL TRANSDUCTION. MAY ACT AS A MEDIATOR OF INSIDE-OUT
CC INTEGRIN SIGNALING. FOCAL ADHESION PROTEIN PART OF THE COMPLEX
CC ILK-PINCH. THIS COMPLEX IS CONSIDERED TO BE ONE OF THE CONVERGENCE
CC POINTS OF INTEGRIN- AND GROWTH FACTOR-SIGNALING PATHWAY. COULD BE
CC IMPLICATED IN MEDIATING CELL ARCHITECTURE, ADHESION TO INTEGRIN
CC SUBSTRATES AND ANCHORAGE-DEPENDENT GROWTH IN EPITHELIAL CELLS.
CC PHOSPHORYLATES BETA-1 AND BETA-3 INTEGRIN SUBUNIT ON SERINE AND
CC THREONINE RESIDUES. BUT ALSO AKT1 AND GSK3B (BY SIMILARITY).
CC -! SUBUNIT: INTERACTS WITH CYTOPLASMIC DOMAIN OF BETA 1 SUBUNIT OF
CC INTEGRIN. COULD ALSO INTERACTS WITH BETA 2, BETA 3 AND/OR BETA 5
CC SUBUNIT OF INTEGRIN. AND ALSO WITH PINCH (BY SIMILARITY).
CC -! SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -! TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LUNG, HEART, KIDNEY,
CC LIVER, BRAIN, SPLEEN AND SKELETAL MUSCLE. WEAKLY EXPRESSED IN
CC TESTIS.
CC -! DOMAIN: A PH-LIKE DOMAIN IS INVOLVED IN PHOSPHATIDYLINOSITOL
CC PHOSPHATE BINDING (BY SIMILARITY).
CC -! PTM: AUTOPHOSPHORYLATED ON SERINE (BY SIMILARITY).
CC -! SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -! SIMILARITY: CONTAINS 3 ANK REPEATS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; U94479; AAB94646.1;
CC HSPSP; Q00421; IAWC.
CC MGD; MGI:1195267; Ilk.
CC InterPro; IPR000719;
CC InterPro; IPR002110;
CC Pfam; PF00023; ank; 3.
CC Pfam; PF00069; pkinase; 2.
CC PROSITE; PS50297; ANK_REPEAT_REGION; 1.
CC PROSITE; PS50088; ANK_REPEAT; 3.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC TransFaser; Serine/threonine-protein kinase; ATP-binding;
CC Phosphorylation; Repeat; ANK repeat.
CC REPEAT 33 62
CC FT REPEAT 66 95
CC FT REPEAT 99 128
CC FT REPEAT 180 212
CC FT DOMAIN 193 446
CC FT BIND 199 207
CC FT BINDING 220 220
CC FT ATP (POTENTIAL).
CC FT ATP (POTENTIAL).

```

Search completed: August 13, 2001, 07:44:52
Job time: 29 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2001, 07:47:24 ; Search time 32.54 Seconds
(without alignments)
866.041 Million cell updates/sec

Title: US-09-509-775-2_COPY_14_226

Perfect score: 1093

Sequence: 1 AYSGLLEELKESILADKSLA.....TPLQVAKGGLILKRMVEG 213

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_16:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_unclassified:*
- 13: sp_vertebrate:*
- 14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1093	100.0	226	4	075832
2	1037	94.9	231	11	Q922X2
3	1035	94.7	231	11	Q922X3
4	511.5	46.8	118	4	095533
5	338	30.9	247	10	Q92Q79
6	312.5	28.6	1943	11	Q61307
7	310	28.4	237	3	Q9P3P6
8	310	28.4	1088	4	Q13484
9	310	28.4	4377	4	Q12955
10	306	28.0	2622	11	Q70511
11	304.5	27.9	636	5	Q9T267
12	301	27.5	1762	11	Q88521
13	297	27.2	1159	5	Q9NCP8
14	297	27.2	2443	5	Q9VSA2
15	290.5	26.6	1136	6	Q9N180
16	288.5	26.4	1786	5	Q17344
17	288.5	26.4	1809	5	Q17487
18	288.5	26.4	1815	5	Q17488
19	288.5	26.4	1867	5	Q17486

20	288.5	26.4	2039	5	Q17489
21	288.5	26.4	6994	5	Q17343
22	287.5	26.3	1848	11	Q61302
23	286.5	26.2	1719	4	Q13768
24	286.5	26.2	1856	4	Q99407
25	286	26.2	843	11	P97582
26	283.5	25.9	1059	4	Q15084
27	283.5	25.9	1166	4	Q9H2K2
28	283.5	25.9	1265	4	Q9HASA
29	282	25.8	1098	11	Q61304
30	281.5	25.8	1327	4	Q95271
31	281	25.7	1181	5	Q9XZ37
32	281	25.7	1181	5	Q9VBP3
33	276	25.3	1549	5	Q24241
34	276	25.3	1549	5	Q9V4B1
35	269.5	24.7	815	6	Q9TV77
36	269.5	24.7	1030	4	Q14974
37	268.5	24.6	658	11	Q62937
38	268.5	24.6	976	11	Q10728
39	267	24.4	963	13	Q90624
40	267	24.4	1004	13	Q90623
41	266	24.3	768	4	Q9ULJ7
42	260.5	23.8	1062	11	Q89019
43	260	23.8	1430	11	Q9JJF7
44	259.5	23.7	1395	5	Q9XZC0
45	258.5	23.7	2119	5	Q9VAU5

ALIGNMENTS

RESULT 1

075832 PRELIMINARY; PRT; 226 AA.

AC 075832; 01-NOV-1998 (TRENBLrel. 08, Created)

DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)

DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)

DE 26S PROTEASOME SUBUNIT p28,

GN DJ889N15.2 OR HUMAN GANKYRIN.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98382579; PubMed=9714768;

RA Hori T., Kato S., Saeki M., DeMartino G.N., Slaughter C.A.,

RA Takeuchi J., Toh-E A., Tanaka K.;

RT "cDNA cloning and functional analysis of p28 (Nas6p) and p40.5

RT (Nas7p), two novel regulatory subunits of the 26S proteasome.";

RL Gene 216:113-122(1998).

RP [2]

RN SEQUENCE FROM N.A.

RP Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

RA Grafham D.;

RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=PLACENTA;

RT Higashitsuji H., Fujita J.;

RT "Enhanced expression of a novel tumour marker in the human

RT hepatomas.";

RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB009619; BAA33215.1; -

DR EMBL; AL031177; CAA20117.1; -

DR EMBL; D83197; BAA34594.1; -

DR HSSP; P42773; 1IHB.

DR InterPro; IPR002110; -

DR Pfam; PF00023; ank; 5.

DR PROSITE; PS50088; ANK_REPEAT; 5.

DR PROSITE; PS50297; ANK_REPEAT; 1.

DR SMART; SM00248; ANK; 1.

KW Proteasome.

```
SQ SEQUENCE 226 AA; 24428 MW; 57158E33146EC7C8 CRC64;

Query Match 100.0%; Score 1093; DB 4; Length 226;
Best Local Similarity 100.0%; Pred. No. 6.7e-91;
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYSKLEELKESILADKSLATRTDQDSRTALHWACSGAGHTEIVEFLLQLGVPVNDKDDAG 60
DB 14 AYSKLEELKESILADKSLATRTDQDSRTALHWACSGAGHTEIVEFLLQLGVPVNDKDDAG 73
QY 61 WSPHLHAASAGRDEIVKALLGKGAQVNAVQNGCTPLHYAASKNRHEIAVMLEGGANPD 120
DB 74 WSPHLHAASAGRDEIVKALLGKGAQVNAVQNGCTPLHYAASKNRHEIAVMLEGGANPD 133
QY 121 AKDHYEATAMHRAAAGNLMKMIHLLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVS 180
DB 134 AKDHYEATAMHRAAAGNLMKMIHLLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVS 193
QY 181 QGASITYENKEEKTPLQVAKGGLGLILKRWVEG 213
DB 194 QGASITYENKEEKTPLQVAKGGLGLILKRWVEG 226

RESULT 2
ID Q922X2 PRELIMINARY; PRT; 231 AA.
AC Q922X2;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE GANKYRIN.
GN PSMD10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC TISSUE=PLACENTA;
RA Higashitsuji H., Fujita J.;
RT "Cloning of mouse gankyrin containing ankyrin repeats.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022022; BAA36969.1; -.
DR HSSP; P42773; 1IHB.
DR MGP; MGI-1858898; Psmid10.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank; 5.
DR PROSITE; PS50088; ANK_REPEAT; 5.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 231 AA; 25115 MW; 5A3DB7027B06594E6 CRC64;

Query Match 94.9%; Score 1037; DB 11; Length 231;
Best Local Similarity 93.9%; Pred. No. 7.7e-86;
Matches 199; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 AYSKLEELKESILADKSLATRTDQDSRTALHWACSGAGHTEIVEFLLQLGVPVNDKDDAG 60
DB 14 AYSKLEELKESILADKSLATRTDQDSRTALHWACSGAGHTEIVEFLLQLGVPVNDKDDAG 73
QY 61 WSPHLHAASAGRDEIVKALLGKGAQVNAVQNGCTPLHYAASKNRHEIAVMLEGGANPD 120
DB 74 WSPHLHAASAGRDEIVKALLGKGAQVNAVQNGCTPLHYAASKNRHEIAVMLEGGANPD 133
QY 121 AKDHYEATAMHRAAAGNLMKMIHLLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVS 180
DB 134 AKDHYEATAMHRAAAGNLMKMIHLLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVS 193
QY 181 QGASITYENKEEKTPLQVAKGGLGLILKRWVE 212
DB 194 QGASITYENKEEKTPLQVAKGGLGLILKRWVE 225
```

```
RESULT 3
Q922X3 PRELIMINARY; PRT; 231 AA.
ID Q922X3;
AC Q922X3;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE GANKYRIN HOMOLOGUE.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RC TISSUE=PLACENTA;
RA Higashitsuji H., Fujita J.;
RT "Cloning of rat gankyrin homologue containing ankyrin repeats.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022014; BAA36954.1; -.
DR HSSP; P42773; 1IHB.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank; 5.
DR PROSITE; PS50088; ANK_REPEAT; 5.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 231 AA; 24985 MW; F5241DC9A816066E CRC64;

Query Match 94.7%; Score 1035; DB 11; Length 231;
Best Local Similarity 94.3%; Pred. No. 1.2e-85;
Matches 200; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

QY 1 AYSKLEELKESILADKSLATRTDQDSRTALHWACSGAGHTEIVEFLLQLGVPVNDKDDAG 60
DB 14 AYSKLEELKESILADKSLATRTDQDSRTALHWACSGAGHTEIVEFLLQLGVPVNDKDDAG 73
QY 61 WSPHLHAASAGRDEIVKALLGKGAQVNAVQNGCTPLHYAASKNRHEIAVMLEGGANPD 120
DB 74 WSPHLHAASAGRDEIVKALLGKGAQVNAVQNGCTPLHYAASKNRHEIAVMLEGGANPD 133
QY 121 AKDHYEATAMHRAAAGNLMKMIHLLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVS 180
DB 134 AKDHYEATAMHRAAAGNLMKMIHLLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVS 193
QY 181 QGASITYENKEEKTPLQVAKGGLGLILKRWVE 212
DB 194 QGASITYENKEEKTPLQVAKGGLGLILKRWVE 225

RESULT 4
Q95533 PRELIMINARY; PRT; 118 AA.
ID Q95533;
AC Q95533;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE DJ889N15.2.2 (26S PROTEASOME SUBUNIT P28 (ANKYRIN REPEAT PROTEIN))
DE (PUTATIVE PARTIAL ISOFORM 2) (FRAGMENT).
GN DJ889N15.2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC TISSUE=PLACENTA;
RA Grafham D.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL031177; CAA20118.1; -.
DR HSSP; Q13625; IYCS.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank; 2.
```


KW Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
KW Multigene family.
FT DOMAIN 1 826 ANION EXCHANGE PROTEIN BINDING DOMAIN.
FT DOMAIN 23 810 24 X ANK MOTIF REPEATS.
FT DOMAIN 856 1437 SPECTRIN BINDING DOMAIN.
FT DOMAIN 1438 1943 REGULATORY DOMAIN (REGULATES BINDING OF ANKYRIN TO SPECTRIN AND THE BAND 3 PROTEIN).
FT REPEAT 23 55 ANK MOTIF 1.
FT REPEAT 56 88 ANK MOTIF 2.
FT REPEAT 89 121 ANK MOTIF 3.
FT REPEAT 122 154 ANK MOTIF 4.
FT REPEAT 155 183 ANK MOTIF 5.
FT REPEAT 184 216 ANK MOTIF 6.
FT REPEAT 217 249 ANK MOTIF 7.
FT REPEAT 250 282 ANK MOTIF 8.
FT REPEAT 283 315 ANK MOTIF 9.
FT REPEAT 316 348 ANK MOTIF 10.
FT REPEAT 349 381 ANK MOTIF 11.
FT REPEAT 382 414 ANK MOTIF 12.
FT REPEAT 415 447 ANK MOTIF 13.
FT REPEAT 448 480 ANK MOTIF 14.
FT REPEAT 481 513 ANK MOTIF 15.
FT REPEAT 514 546 ANK MOTIF 16.
FT REPEAT 547 579 ANK MOTIF 17.
FT REPEAT 580 612 ANK MOTIF 18.
FT REPEAT 613 645 ANK MOTIF 19.
FT REPEAT 646 678 ANK MOTIF 20.
FT REPEAT 679 711 ANK MOTIF 21.
FT REPEAT 712 744 ANK MOTIF 22.
FT REPEAT 745 777 ANK MOTIF 23.
FT REPEAT 778 810 ANK MOTIF 24.
FT VARSPLIC 1 805 MISSING (IN ISOFORM 7 AND ISOFORM 8).
FT VARSPLIC 1 849 MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT VARSPLIC 833 853 MISSING (IN ISOFORM 7 AND ISOFORM 8).
FT VARSPLIC 834 855 VRKASPEKLSGDGEYISDGEEG -> GDKCTWFKPKVQEV LVK (IN ISOFORM 5).
FT VARSPLIC 850 855 SDGEEG -> MALPHS (IN ISOFORM 2 AND ISOFORM 3).
FT VARSPLIC 852 852 G -> GDKCTWFKPKVQEVLYKS (IN ISOFORM 4 AND ISOFORM 6).
FT VARSPLIC 1569 1764 MISSING (IN ISOFORM 3, ISOFORM 4 AND ISOFORM 7).
SQ SEQUENCE 1943 AA; 212013 MW; B23C4B6F82904804 CRC64;

Query Match 28.6%; Score 312.5; DB 11; Length 1943;
Best Local Similarity 35.5%; Pred. No. 2.9e-19;
Matches 72; Conservative 31; Mismatches 67; Indels 33; Gaps 1;

QY 29 TALHWACSAGHTEIVEFLLQLGVVNDKDDAGWSPLHIAASAGRDEIVKALLGK----- 82
DB 517 TPLHIAAREGHEDVNAFLLDHAGLSITTKGFTPLHVAAYKGLVAVSLLLOKSASPD 576
QY 83 -----GAQVNAVNGCTGPLHYAASKNRHEIAVMLEGG 115
DB 577 AGKSGLTPLHVAAYHNDQKVALLLDQASPHAAAKNGYTPHIAAKKNQMDIATSLLEY 636
QY 116 GANPDADHYEATAMHRAAAGNLKMIHILLYKASTNIQTEGNTPLHLACDEERVEEA 175
DB 637 GADANAVTQGTASVLAHQEGHVDMSLLSRNANVNLNKSGLTPLHLAAQEDRVNVA 696
QY 176 KLVVSGASIVTENKEETPLQV 198
DB 697 EVLVNQGAHVDAQTRMGYTPLVH 719

RESULT 7
QY9P3P6 PRELIMINARY; PRT; 237 AA.
AC QY9P3P6;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE RELATED TO 26S PROTEASOME SUBUNIT P28.
GN B7F18.30.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL398981; CAB97304.1;
DR InterPro; IPR002110;
DR Pfam; PF00023; ank; 6.
DR PROSITE; PS00088; ANK_REPEAT; 3.
DR PROSITE; PS0297; ANK_REP_REGION; 1.
DR SMART; SM00248; ANK; 1.
KW Proteasome. 237 AA; 25995 MW; 6E2B6EFC69D909B1 CRC64;
SQ SEQUENCE 237 AA; 25995 MW; 6E2B6EFC69D909B1 CRC64;

Query Match 28.4%; Score 310; DB 3; Length 237;
Best Local Similarity 34.5%; Pred. No. 2.9e-20;
Matches 78; Conservative 30; Mismatches 90; Indels 28; Gaps 4;

QY 1 AYSGKLEELKESLADKSLATRTDQSRALHWACSAGHTEIVEFLL-QLGVVNDKDDA 59
DB 12 ARDGKASIVESLLNANPKLAQRKDDGRLPIHWACSYNRKVEVLVNGKDFDPVEDDM 71
QY 60 GWSPLHIAASA-GRDEIVKALLGGAQVNAVNGCTPLHYAASKNRHEIAVMLEGGAN 118
DB 72 GMTFFMISASVSDSDAIDLLSRGADINQTHQSOTALHFTASKNNIDLARKLLSPDMK 131
QY 119 PD-----AKDHVEATAMHRAAAGNLKMIHILLYKASTNIQTEGNTPLHLACDEERVE 173
DB 132 PKPASVRVNDKRGQYPLHRAAAGISVPMINLLQHKSPINASDNAGYTPHHLHVAEGHGH 191
QY 174 EAKLVVSGASI-----YIENKEETKPTLOV 198
DB 192 AAVALLKAGAEATDKDMOGYLALDLAPDKDVRFRFIEKEAKEGIEL 237

RESULT 8
QY13484 PRELIMINARY; PRT; 1088 AA.
AC QY13484;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE ANKYRIN G119.
GN ANK3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RX MEDLINE=96234670; Pubmed=8666667;
RA Kashgarian M., Morrow J.S., Ardito T., Stabach P.R., Mann A.S.,
RA Devarajan P.;
RT "Identification of a small cytoplasmic ankyrin (AnkG119) in the kidney and muscle that binds beta I sigma spectrin and associates with the Golgi apparatus";
RL J. Cell Biol. 133:819-830(1996).
DR EMBL; U43965; AAB08437.1;
DR HSP; P55273; IBD8.
DR InterPro; IPR000906;

DR InterPro: IPR002110; -.
DR Pfam: PF00023; ank; 12.
DR Pfam: PF00791; ZU5; 1.
DR PROSITE: PS50088; ANK_REPEAT; 12.
DR PROSITE: PS50297; ANK_REP_REGION; 1.
DR SMART: SM00218; ZU5; 1.
SQ SEQUENCE 1088 AA; 119427 MW; 769C88040A78DE86 CRC64;

Query Match 28.4%; Score 310; DB 4; Length 1088;
Best Local Similarity 37.0%; Pred. No. 2.2e-19;
Matches 71; Conservative 36; Mismatches 85; Indels 0; Gaps 0;

QY 7 EELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFLLQLGVVNDKDDAGWSPPLHI 66
DB 166 EDVAAPFLLDHGASLSITTKGFTPLHVAAYKGLVAVANLLQKSAPDAAGKSLTPLHV 225
QY 67 AASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHEIAVMLEGANPDARDHYE 126
DB 226 AAHYDNQKVALLLDQGASPHAAAKNGYTPPLHIAAKKNQMDIATTLLEYGADANAVTROG 285
QY 127 ATAMHRAAAGNLMKIHILLYKASTNIODTEGNTPLHLACDEEREEAKLLVSGQASII 186
DB 286 IASVHLAAQEGHVMVSLLLGRNANVNLNKSGLTPLHLAAQEDRVNVAEVLVNOGAHVD 345
QY 187 IENKEEKTPLQV 198
DB 346 AQTCKMGYTPPLHV 357

RESULT 9
Q12955 PRELIMINARY; PRT; 4377 AA.
AC Q12955;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-MAR-2001 (TremBLrel. 16, Last annotation update)
DE ANKYRIN G.
GN ANK-3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN STEM;
RX MEDLINE=95138209; PubMed=7836469;
RA Kordeli E., Lambert S., Bennett V.;
RT "AnkyrinG, A new ankyrin gene with neural-specific isoforms localized at the axonal initial segment and node of Ranvier.";
RL J. Biol. Chem. 270:2352-2359(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN STEM;
RA Carpenter S.S.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL: U13616; AA64834.1; -.
DR HSSP: P55273; 1B08.
DR InterPro: IPR000488; -.
DR InterPro: IPR000906; -.
DR InterPro: IPR002110; -.
DR Pfam: PF00023; ank; 23.
DR Pfam: PF00531; death; 1.
DR Pfam: PF00791; ZU5; 1.
DR PROSITE: PS50088; ANK_REPEAT; 21.
DR PROSITE: PS50297; ANK_REP_REGION; 1.
DR PROSITE: PS50017; DEATH_DOMAIN; 1.
DR SMART: SM00005; DEATH; 1.
SQ SEQUENCE 4377 AA; 480399 MW; F42379E55768B684 CRC64;

Query Match 28.4%; Score 310; DB 4; Length 4377;
Best Local Similarity 37.0%; Pred. No. 1.4e-18;

Matches 71; Conservative 36; Mismatches 85; Indels 0; Gaps 0;

QY 7 EELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFLLQLGVVNDKDDAGWSPPLHI 66
DB 545 EDVAAPFLLDHGASLSITTKGFTPLHVAAYKGLVAVANLLQKSAPDAAGKSLTPLHV 604
QY 67 AASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHEIAVMLEGANPDARDHYE 126
DB 605 AAHYDNQKVALLLDQGASPHAAAKNGYTPPLHIAAKKNQMDIATTLLEYGADANAVTROG 664
QY 127 ATAMHRAAAGNLMKIHILLYKASTNIODTEGNTPLHLACDEEREEAKLLVSGQASII 186
DB 665 IASVHLAAQEGHVMVSLLLGRNANVNLNKSGLTPLHLAAQEDRVNVAEVLVNOGAHVD 724
QY 187 IENKEEKTPLQV 198
DB 725 AQTCKMGYTPPLHV 736

RESULT 10
O70511 PRELIMINARY; PRT; 2622 AA.
AC O70511;
DT 01-AUG-1998 (TremBLrel. 07, Created)
DT 01-AUG-1999 (TremBLrel. 11, Last sequence update)
DT 01-MAR-2001 (TremBLrel. 16, Last annotation update)
DE 270 KDA ANKYRIN G ISOFORM (ANKYRING) (FRAGMENT).
GN ANK3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98417633; PubMed=9744885;
RA Zhang X., Bennett V.;
RT "Restriction of 480/270-kD ankyrin G to axon proximal segments requires multiple ankyrin G-specific domains.";
RL J. Cell Biol. 142:1571-1581(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Carpenter S.S.; Zhang X.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 934-1220 FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-SKELETAL MUSCLE;
RA Kordeli E., Ludosky M.A., Deprette C., Frappier T., Cartaud J.;
RL J. Cell Sci. 0:0-0(1998).
DR EMBL: AF102552; AAC78143.1; -.
DR EMBL: AF065150; AAC18853.1; -.
DR HSSP: P55273; 1B08.
DR InterPro: IPR000488; -.
DR InterPro: IPR000906; -.
DR InterPro: IPR002110; -.
DR Pfam: PF00023; ank; 24.
DR Pfam: PF00531; death; 1.
DR Pfam: PF00791; ZU5; 1.
DR PROSITE: PS50088; ANK_REPEAT; 21.
DR PROSITE: PS50297; ANK_REP_REGION; 1.
DR PROSITE: PS50017; DEATH_DOMAIN; 1.
DR SMART: SM00248; ANK; 1.
DR SMART: SM00005; DEATH; 1.
DR SMART: SM00218; ZU5; 1.
DR NON_TER 2622 2622
SQ SEQUENCE 2622 AA; 284456 MW; 67B34830D3AC884E CRC64;

Query Match 28.0%; Score 306; DB 11; Length 2622;
Best Local Similarity 36.5%; Pred. No. 1.7e-18;
Matches 70; Conservative 36; Mismatches 86; Indels 0; Gaps 0;

QY 7 EELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFLLQLGVVNDKDDAGWSPPLHI 66
DB 166 EDVAAPFLLDHGASLSITTKGFTPLHVAAYKGLVAVANLLQKSAPDAAGKSLTPLHV 225
QY 67 AASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHEIAVMLEGANPDARDHYE 126
DB 226 AAHYDNQKVALLLDQGASPHAAAKNGYTPPLHIAAKKNQMDIATTLLEYGADANAVTROG 285
QY 127 ATAMHRAAAGNLMKIHILLYKASTNIODTEGNTPLHLACDEEREEAKLLVSGQASII 186
DB 286 IASVHLAAQEGHVMVSLLLGRNANVNLNKSGLTPLHLAAQEDRVNVAEVLVNOGAHVD 345
QY 187 IENKEEKTPLQV 198
DB 346 AQTCKMGYTPPLHV 357

Db 553 EDVAEFLDHGASLSITTKKGFPLHVAAYKGLVLEASLLOKSASPDAAAGKSLTPLHV 612
 QY 67 AASAGRDEIVKALGKAQVNAVONQCTPLHYAASKNRHEIAVMLEGGANPDADKHYE 126
 Db 613 AAHYDNOKVALLLLDOGASPHAAKNGYTPHLHAAKKNQMDIATSLLEYGADANPYTRQG 672
 QY 127 ATAMHRAAAKGNLMIHLLYYKASTNIQDTGNTPLHACDEERVEEAKLLYSQASIV 186
 Db 673 IASVHLAAQGGHVDMSLLSRANVNLNKSGLTPLHLAAQEDRVNVAEVLVNOGAHVD 732
 QY 187 IENKEEKTPLQV 198
 Db 733 AQTRMGYTPLVH 744
 RESULT 11
 Q9TZ67 PRELIMINARY; PRT: 636 AA.
 AC Q9TZ67;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE F40G9.1 PROTEIN.
 GN F40G9.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showken R.,
 RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RL elegans.";
 RN Nature 368:32-38(1994).
 RP [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF099919; AAC68798.1; -;
 DR HSP; Q00421; IAWC.
 DR InterPro: IPR002110; -;
 DR Pfam; PF00023; ank; 5.
 DR PROSITE; PS50088; ANK_REPEAT; 4.
 DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
 DR SMART; SM00248; ANK; 1.
 SQ SEQUENCE 636 AA: 73229 MW; AB55162AF5D5B1C0 CRC64;
 Query Match 27.9%; Score 304.5; DB 5; Length 636;
 Best Local Similarity 34.6%; Pred. No. 3.4e-19;
 Matches 85; Conservative 39; Mismatches 75; Indels 47; Gaps 7;
 QY 1 AYSGKLEELKESILADKSLATRTDQDSRTAL-----HWACSAGH 39
 Db 379 AVGGSPLQFAINDPEAHKTD-DVRVGMWNLTKIENLFFNRKICDFRKNYSKNTKN 437

QY 40 TEIVFELLO-----LGVPVNDK-----DAGWSPHLIAASAGRDEIVK 77
 Db 438 PFSEFFETIDFEHENSQNALKKFSLKIDFDLNCLLPLGWTPLMIASSAGRVDVVR 497
 QY 78 ALIG-KGAQVNAVONQCTPLHYAASKNRHEIAVMLEGGAN-PDAKDHYEATAMHRAAA 135
 Db 498 YLLTLPDVPDKTNSNKQTSLSHYACSKNHEIVKLLIEADPNILNLPDFEGATALHRAAS 557
 QY 136 KGNLMIHLLYY-KASTNIQDTGNTPLHACDEERVEEAKLLYSQASIVYENKEET 194
 Db 558 RGNQDVTIVRALVSTGKSLDRQDGEQNTALHLACDENRGDVAILLVNRGADMKMLNKEQT 617
 QY 195 PLQVAK 200
 Db 618 PLEMLK 623
 RESULT 12
 O88521 PRELIMINARY; PRT: 1762 AA.
 ID O88521;
 AC O88521;
 DT 01-NOV-1998 (Tremblrel. 08, Created)
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE 190 KDA ANKYRIN ISOFORM.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus;
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=98395113; PubMed=9727010;
 RA Thevananthar S., Kollu A.H., Devatajan P.;
 RT Identification of a novel ankyrin isoform (AnkG190) in kidney and
 RT lung that associates with the plasma membrane and binds alpha-Na, k-
 RT ATPase.";
 RL J. Biol. Chem. 273:23952-23958(1998).
 DR EMBL; AF069525; AAC34809.1; -;
 DR HSP; P55273; 1B18.
 DR InterPro: IPR000169; -;
 DR InterPro: IPR000488; -;
 DR InterPro: IPR000906; -;
 DR InterPro: IPR002110; -;
 DR Pfam; PF00023; ank; 24.
 DR Pfam; PF00531; death; 1.
 DR Pfam; PF00791; ZU5; 1.
 DR PROSITE; PS50088; ANK_REPEAT; 20.
 DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
 DR PROSITE; PS50017; DEATH_DOMAIN; 1.
 DR PROSITE; PS00139; THIOLEPROTEASE_CYS; UNKNOWN_1.
 DR SMART; SM00005; DEATH; 1.
 SQ SEQUENCE 1762 AA; 191898 MW; 9023280086A7BF4E CRC64;
 Query Match 27.5%; Score 301; DB 11; Length 1762;
 Best Local Similarity 35.9%; Pred. No. 2.8e-18;
 Matches 69; Conservative 36; Mismatches 87; Indels 0; Gaps 0;
 QY 7 EELKESILADKSLATRTDQDSRTALHWACSAGHPEIVEFLIQLGVPVNDKDDAGWSPHLI 66
 Db 539 EDVAEFLDHGAFSLITTKKGFPLHVAAYKGLVLEASLLOKSASPDAAAGKSLTPLHV 598
 QY 67 AASAGRDEIVKALGKAQVNAVONQCTPLHYAASKNRHEIAVMLEGGANPDADKHYE 126
 Db 599 AAHYDNOKVALLLLDOGASPHAAKNGYTPHLHAAKKNQMDIATSLLEYGADANAVTRQG 658
 QY 127 ATAMHRAAAKGNLMIHLLYYKASTNIQDTGNTPLHACDEERVEEAKLLYSQASIV 186
 Db 659 IASVHLAAQGGHVDMSLLSRANVNLNKSGLTPLHLAAQEDRVNVAEVLVNOGAHVD 718
 QY 187 IENKEEKTPLQV 198

```
Db 719 AOTRMGYTPLHV 730
RESULT 13
Q9NCP8 PRELIMINARY; PRT; 1159 AA.
AC Q9NCP8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE ANKYNR 2.
GN ANK2 OR CG7462.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20304926; PubMed=10844021;
RA Bouley M., Tian M.-Z., Paisley K., Shen Y.-C., Malhotra J.D.,
RA Hortsch M.;
RT "The li-type cell adhesion molecule neuroglian influences the
RT stability of neural ankyrin in the Drosophila embryo but not its
RT axonal localization.";
RL J. Neurosci. 20:4515-4523(2000).
DR EMBL; AF190635; AAF73309.1; -
DR FlyBase; FBgn0017645; Ank2.
DR InterPro; IPR000906; -.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00791; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 21.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 1159 AA; 125769 MW; 9ED146C9E483DE58 CRC64;

Query Match 27.2%; Score 297; DB 5; Length 1159;
Best Local Similarity 29.7%; Pred. No. 3.6e-18;
Matches 71; Conservative 43; Mismatches 87; Indels 38; Gaps 4;

Qy 12 SLADK--SLATRTDQDSRTALHWACSAAGHTEIVEFLLQLGVPVNDKDDAGWSP----- 63
Db 255 SLLLEKGGNIEAKT-RDGLTPLHCAARSGHQVVDMLLGERGAPISAKTKNGLAPLHMAAQ 313
Qy 64 -----LHIAASAGRDEIVKALLGKGAQVNAVNGCTP 96
Db 314 GEHVDAARILLYHRAPVDEVTVDTALTALHVAACHGVVRVAKLLLDNRNADANARALNGFTP 373
Qy 97 LHYAASKNRHETAVMLLEGGANPDADKHYEATAMHRAAKGNLKMHIHLLYKASTNIQD 156
Db 374 LHIAACKNRKLVVLELLRHGASISATTESGLTPLHVAAFMGCMNIVLYLQHDASPDVPT 433
Qy 157 TEGNTPPLHACDEEREEAKILVSOGASIIYENKEETPLQVAK--GGIGLILKRWVSG 213
Db 434 VRGETPLHARANGTDIIRILLRNGAQVDARAREQOTPLHTASRLGNVDIVMLLQHG 492

RESULT 14
Q9VSA2 PRELIMINARY; PRT; 2443 AA.
AC Q9VSA2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE ESTS:149B105 PROTEIN.
GN ANK2 OR ESTS:149B105 OR CG7462.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;

RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.N., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.B., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector A., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003558; AAF50525.1; -.
DR HSSP; P42773; 1IHB.
DR FlyBase; FBgn0017645; Ank2.
DR InterPro; IPR000488; -.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 2443 AA; 269538 MW; 7561DC5CF56812DC CRC64;

Query Match 27.2%; Score 297; DB 5; Length 2443;
Best Local Similarity 29.7%; Pred. No. 9.8e-18;
Matches 71; Conservative 43; Mismatches 87; Indels 38; Gaps 4;

Qy 12 SLADK--SLATRTDQDSRTALHWACSAAGHTEIVEFLLQLGVPVNDKDDAGWSP----- 63
Db 255 SLLLEKGGNIEAKT-RDGLTPLHCAARSGHQVVDMLLGERGAPISAKTKNGLAPLHMAAQ 313
Qy 64 -----LHIAASAGRDEIVKALLGKGAQVNAVNGCTP 96
Db 314 GEHVDAARILLYHRAPVDEVTVDTALTALHVAACHGVVRVAKLLLDNRNADANARALNGFTP 373
Qy 97 LHYAASKNRHETAVMLLEGGANPDADKHYEATAMHRAAKGNLKMHIHLLYKASTNIQD 156
Db 374 LHIAACKNRKLVVLELLRHGASISATTESGLTPLHVAAFMGCMNIVLYLQHDASPDVPT 433
Qy 157 TEGNTPPLHACDEEREEAKILVSOGASIIYENKEETPLQVAK--GGIGLILKRWVSG 213
Db 434 VRGETPLHARANGTDIIRILLRNGAQVDARAREQOTPLHTASRLGNVDIVMLLQHG 492
```

Search completed: August 13, 2001, 07:47:25
Job time: 181 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: August 13, 2001, 07:45:23 ; Search time 1109.56 Seconds
(without alignments)
10873.539 Million cell updates/sec

Title: US-09-509-775-1
Perfect score: 780
Sequence: 1 tggtagactctaacgctg.....gaatggtggaaggttaaca 780

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_bal: *
2: gb_ba2: *
3: gb_ba3: *
4: gb_in1: *
5: gb_in2: *
6: gb_in3: *
7: gb_om: *
8: gb_ov: *
9: gb_pat1: *
10: gb_pat2: *
11: gb_ph: *
12: gb_pl1: *
13: gb_pl2: *
14: gb_pl3: *
15: gb_pl4: *
16: em_bal: *
17: em_ba2: *
18: em_fun: *
19: em_htgo_hum: *
20: em_htgo_inv: *
21: em_htgo_rod: *
22: em_htg_hum1: *
23: em_htg_hum2: *
24: em_htg_hum3: *
25: em_htg_hum4: *
26: em_htg_hum5: *
27: em_htg_hum6: *
28: em_htg_hum7: *
29: em_htg_hum8: *
30: em_htg_inv1: *
31: em_htg_inv2: *
32: em_htg_other: *
33: em_htg_rod: *
34: em_hum1: *
35: em_hum2: *
36: em_hum3: *
37: em_hum4: *
38: em_hum5: *
39: em_hum6: *
40: em_hum7: *
41: em_in: *
42: em_om: *
43: em_or: *

44: em_ov: *
45: em_pat: *
46: em_ph: *
47: em_pl: *
48: em_ro: *
49: em_sts: *
50: em_sy: *
51: em_un: *
52: em_vi: *
53: gb_sts1: *
54: gb_sts2: *
55: gb_sts3: *
56: gb_sy: *
57: gb_un: *
58: gb_vil: *
59: gb_vil2: *
60: gb_htg1: *
61: gb_htg2: *
62: gb_htg3: *
63: gb_htg4: *
64: gb_htg5: *
65: gb_htg6: *
66: gb_htg7: *
67: gb_htg8: *
68: gb_htg9: *
69: gb_htg10: *
70: gb_htg11: *
71: gb_htg12: *
72: gb_htg13: *
73: gb_htg14: *
74: gb_htg15: *
75: gb_htg16: *
76: gb_htg17: *
77: gb_htg18: *
78: gb_htg19: *
79: gb_htg20: *
80: gb_htg21: *
81: gb_htg22: *
82: gb_htg23: *
83: gb_htg24: *
84: gb_htg25: *
85: gb_pr1: *
86: gb_pr2: *
87: gb_pr3: *
88: gb_pr4: *
89: gb_pr5: *
90: gb_pr6: *
91: gb_pr7: *
92: gb_pr8: *
93: gb_pr9: *
94: gb_rol: *
95: gb_rol2: *
96: gb_in4: *
97: gb_pr10: *
98: em_ba3: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	780	100.0	1544	91 D83197	D83197 Homo sapien
2	706	90.5	1468	10 E12804	E12804 cDNA encodi
3	706	90.5	1468	85 AB009619	AB009619 Homo sapi
4	593	76.0	696	94 AB022014	AB022014 Rattus no
5	584	74.9	696	94 AB022022	AB022022 Mus muscu
6	562	72.1	180905	62 AC012149	AC012149 Homo sapi
7	532	68.2	105334	92 HS914P20	AL034553 Human DNA
8	488.6	62.6	175550	61 AC009863	AC009863 Homo sapi

9	398.2	51.1	199882	90	AL354720	Human DNA
10	223.8	28.7	145483	74	AC072018	Homo sapi
c	11	211	27.1	123395	92	HS889N15
	12	211	27.1	178492	78	AL137065
c	13	115.6	14.8	5069	6	DROANKY
	14	115.6	14.8	10152	63	AC014567
c	15	115.6	14.8	114334	5	AE003845
	16	109.4	14.0	4408	85	AB002377
c	17	98.4	12.6	2529	95	RNU65916
	18	96	12.3	6278	93	HSRANK2
c	19	96	12.3	12518	93	HSRANK440
	20	93.4	12.0	5235	9	AR048831
c	21	93.4	12.0	5235	9	AR048832
	22	93.4	12.0	5235	9	AR050426
c	23	93.4	12.0	5235	9	AR050427
	24	93.4	12.0	5503	9	AR048829
c	25	93.4	12.0	5503	9	AR048830
	26	93.4	12.0	5503	9	AR050424
c	27	93.4	12.0	5503	9	AR050425
	28	89.8	11.5	4818	5	AF190635
c	29	89.2	11.4	102552	65	AC019368
	30	89.2	11.4	273744	4	AE003558
c	31	88.4	11.3	3454	97	HSU43965
	32	88.4	11.3	14770	97	HSU13616
c	33	88	11.3	3520	94	MMANK1AC
	34	88	11.3	6238	94	MUSANK1A
c	35	88	11.3	8145	94	MMANK1AA
	36	86.8	11.1	7084	94	MUSANK3B
c	37	86.6	11.1	911	9	AR048821
	38	86.6	11.1	911	9	AR048822
c	39	86.6	11.1	911	9	AR050416
	40	86.6	11.1	911	9	AR050417
c	41	85.6	11.0	909	9	AR048823
	42	85.6	11.0	909	9	AR048824
c	43	85.6	11.0	909	9	AR050418
	44	85.6	11.0	909	9	AR050419
45	85.2	10.9	6192	97	HUMANK	M28880 Human eryth
ALIGNMENTS						
RESULT	1	D83197	1544 bp	mRNA	PRI	06-JAN-1999
LOCUS		D83197				
DEFINITION		Homo sapiens mRNA for ankyrin repeat protein, complete cds.				
ACCESSION		D83197				
VERSION		D83197.1		GI:3893154		
KEYWORDS		human gankyrin; ankyrin repeat protein.				
SOURCE		Homo sapiens female placenta cDNA to mRNA, clone:yf68b03.				
ORGANISM		Homo sapiens				
REFERENCE	1	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1	(bases 1 to 1544)				
		Higashitsuji, H.				
TITLE	1	Direct Submission				
		Submitted (20-JAN-1996) to the DDBJ/EMBL/GenBank databases. Hiroaki				
JOURNAL	1	Higashitsuji, Faculty of Medicine, Kyoto University, Clinical				
		Molecular Biology; Shogoin kawahara-cho 54, Sakyo-ku, Kyoto, Kyoto				
FEATURES	1	605, Japan (E-mail:higashitsuji@viru1.virus.kyoto-u.ac.jp,				
		Tel:81-75-3753, Fax:81-75-3750)				
REFERENCE	2	(sites)				
		Higashitsuji, H. and Fujita, J.				
AUTHORS	1	Enhanced expression of a novel tumour marker in the human hepatomas				
		Unpublished (1996)				
TITLE	1	Location/Qualifiers				
		1..1544				
FEATURES	1	/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
source	1	/clone="yf68b03"				
		/sex="female"				
gene	1	/tissue_type="placenta"				
		99..779				

CDS	/gene="human gankyrin"	99..779
	/gene="human gankyrin"	
	/codon_start=1	
	/product="ankyrin repeat protein"	
	/protein_id="BAA34594.1"	
	/db_xref="GI:3893155"	
	/translation="MEGCVSNDVMCNLAYSKLEELKESILADKSLATRTDSDSTAL	
	HWACSAQTEIVEFLLQIFGVNDDKAGNSPLHIAASAGRDEIVKALLGGAQVNAV	
	NONGCTPLHYAASKNRHEIAVMLLEGGANPDAKHIEATAMRAAKNLKMLHILLY	
	YKASNIODTEGNTPLHLACDEERVEEAKLLVSQASTYIENKKEKTPLOVAKGLGL	
	ILKRNVEG"	
BASE COUNT	435 a	287 c 354 g 468 t
ORIGIN	Query Match	
	Best Local Similarity 100.0%; Score 780; DB 91; Length 1544;	
	Matches 780; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	tgtgtgaagctctaaacggtgttttgaactggcgtagccggagccggcgacgtgagcgggc 60
Db	3	TGCTGAAGCTCTAACGGCTGTTTGACTGGCGTAGCGGCGCGGCGACGTGAGCGGGC 62
QY	61	gttgcctgcgcgcacaaagtgtgtctggggacagcgaataaggaggggtgtgtctaaaccta 120
Db	63	GTTGCTCGCGCGCAACAAGTAGTTGCTGGGACAGCAATGGAGGGGTGTGTCTAAACCTA 122
QY	121	atgctgcgaacctggcctacacgcggagagctggaagagtggaagagagatctctggcc 180
Db	123	ATGGTCTGCAACCTGGCCCTACAGCGGGAAGCTGGAAGAGTTGAAGGAGAGATTTCTGGCC 182
QY	181	gataaatccctgctactagaactgacccaggagcagcagcaactgcattgcactggggcatgc 240
Db	183	GATAAATCCCTGCTACTAGAACTGACACGAGCAGCAGAACTGCATTGCAATGCGCATGC 242
QY	241	tcagctggacatacacgaataattgtgaatttttgcacacttgagtgagtgagtgagtgat 300
Db	243	TCAGCTGGCATACAGAAATTTTGAATTTTTCGTCACCTTGAGTGGCAGTGAATGAT 302
QY	301	aaagacatgcaggttggtctctctctcattctgctgctgctgctgctgctgctgctgctgct 360
Db	303	AAAGACATGCAGGTGGTGTCT 362
QY	361	gtaaaagcccttctggaaaagtgctcgaagtgaatgctgtcgaatcgaatcgaatcgaatcgaat 420
Db	363	GTAAGGCCCTTCTGGAAAAGGTGCTCAAGTGAATGCTGCTCAATCAAAATGCTGTACT 422
QY	421	ccottacattatcagcttcgaaaacagcagcagcagcagcagcagcagcagcagcagcagcagc 480
Db	423	CCCTTACATTATGCAGCTTCGAAAACAGGCATGAGATCGCTGTCAFTGTTTACTTGGAGGC 482
QY	481	ggggctaataccagatgctaaggaccattatgaggtctacagcaatgcacccgggcagcagcc 540
Db	483	GGGGCTAATCCAGATGCTAAGGACCATATATGAGGCTACAGCAATGCACCGGGCAGCAGCC 542
QY	541	aagggttaactggaagtgttcattatctctctctctctctctctctctctctctctctctctct 600
Db	543	AAGGGTAACCTGAAGATGATTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 602
QY	601	gacactgaaggttaacactcctctacacttagcctgtgtaggagagagtggaagaagca 660
Db	603	GACACTGAGGGTACACTCTCTTACACTTACCTTGTGATGAGGAGAGTGGGAAGCA 662
QY	661	aaactgctgtgtcccaagagcaagtatttacctattgagataaagaagaagaagaacccc 720
Db	663	AAACTGCTGTGTGCCAAGGAGCAAGTATTTACATTGAGAAATAAGAGAGAAAGACACCC 722
QY	721	ctgcaagtggcccaaggtggcctgggttttactactcaagagaatgggtggaaggttaaca 780
Db	723	CTGCAAGTGGCCAAAGGTGGCTGGGTTTAACTCAAGAGATGGTGGAGAGGTAAACA 782

YKASTNIQDTGNTPLHLACDERVEEAKLLVSOQASIIYIENKEKTPLOVAKGGLGL
ILKRWVEG"

polyA_signal 1453...1458
BASE COUNT 424 a 268 c 324 g 452 t
ORIGIN

```
Query Match      90.5%; Score 706; DB 85; Length 1468;
Best Local Similarity 100.0%; Pred. No. 6.7e-204;
Matches 706; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 aagtaattgctggagcagcgaataagaggggtgtgtgttaacacctaagtgtcgcaacct 134
DB 1  AAGTAGTTGCTGGGACAGCGAATGGAGGGGTGTGTGCTAACCTAATGGTTCGAACCT 60
QY 135 ggctacagcgggaagctggaagagtgtaaggagagttattctggccgataataatccctggc 194
DB 61  GGCTACAGCGGGAAGCTGGAAGAGTTGAAGGAGAGTATTCCTGGCGGATAAATCCCTGCG 120
QY 195 tactagaactgaccagcagcagcagaactcattgcactggcgatcctcagctcagctgacatac 254
DB 121 TACTAGAACTGACGAGCAGCAGCAACTGCATTGCTACTGGGATGCTCAGCTGGGACATAC 180
QY 255 agaaattgtgaatttttctgcaacttgagtgccagtgaaatgataaagacgatgcagg 314
DB 181 AGAAATGTTGAATTTTGTGCAACTTGGAGTGGCAGTGAATGATAAAGACGATGCAGG 240
QY 315 ttggtctctctcatatgcggctctctgctggcgggagtgagattgtaaaagccctct 374
DB 241 TTGGTCTCTCTTCATATTCGGGCTTCTGCTGGCGGATGAGATTGTAAGGCCCTTCT 300
QY 375 ggcaaaagtgctcaagtaagtctgctcaatcaaaatgctgtactccctacattatgc 434
DB 301 GGGAAAGGTGCTCAAGTGAATGCTGCTCAATCAAAATGGCTGTACTCCCTTACATTTATGC 360
QY 435 agcttggaaaaacagcagcagatgcctgctgctatgttacttgggaaggcggggttaataccaga 494
DB 361 AGCTTGGAAAAACAGCAGCATGAGATGCTGTGATGTTACTTGGAAAGCGGGGCTAATCCAGA 420
QY 495 tctaaggaccattatgagctacacgaatgcacggcgagcagcgaagtgtaacttgaa 554
DB 421 TGTAAGGACCATATGAGGCTACAGCAATGCAACCGGGCAGCAGCCAAAGGTTAACTTGA 480
QY 555 gatgattcatctctctgctactcaaaagcatcccaaaacatccaaagacactgagggtaa 614
DB 481 GATGATTATATCTTCTGTACTTACAAAGCATCCCAAAACATCCAAAGACACTGAGGGTAA 540
QY 615 cactctctacattagcctgtgtatgagagagagtggaagaagcaaaactgctggtgc 674
DB 541 CACTCTCTTACACTTAGCCTGTGATGAGGAGAGTGGGAAGAGCAAAACTGCTGCTGTC 600
QY 675 ccaaggagcgaagtatttacattgagaataaagagaaaagacacccctcaagtggccaa 734
DB 601 CCAAGGAGCAAGTATTTACATTGAGATAAAGAAAGAAAGACACCCCTGCAAGTGCCCAA 660
QY 735 aggtggcctgggtttaaatactcaagaagaatggtgggaaggttaaca 780
DB 661 AGGTGGCTTGGTTTAACTACTCAAGAGAAATGGTGGAAAGGTTAAACA 706
```

RESULT 4
LOCUS AB022014 696 bp mRNA ROD 09-JAN-1999
DEFINITION Rattus norvegicus mRNA for gankyrin homologue, complete cds.
ACCESSION AB022014
VERSION AB022014.1 GI:4127228
KEYWORDS rat gankyrin homologue.
SOURCE Rattus norvegicus adult female placenta cDNA to mRNA.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 696)

AUTHORS
TITLE
JOURNAL

Higashitsuji, H. and Fujita, J.

Direct Submission

Submitted (04-JAN-1999) to the DDBJ/EMBL/GenBank databases. Hiroaki Higashitsuji, Faculty of Medicine, Kyoto University, Clinical Molecular Biology; Shogoin Kawahara-cho 54, Sakyo-ku, Kyoto, Kyoto 606-8507, Japan (E-mail:hhigashi@viru.s.kyoto-u.ac.jp, Tel:81-75-751-3753, Fax:81-75-751-3750)

2 (bases 1 to 696)

Higashitsuji, H. and Fujita, J.

Cloning of rat gankyrin homologue containing ankyrin repeats

Published Only in Database (1999) In press

Location/Qualifiers

source

1. .696
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/dev_stage="adult"
/sex="female"
/tissue_type="placenta"
1. .696
/codon_start=1
/product="gankyrin homologue"
/protein_id="BAA36954.1"
/db_xref="GI:4127228"

/translation="MEGCVNLMVNCNLAYNKLDELKESILADKSLATFRDQDSRTAL
HWACSGAGTEIVEFLQLQGVNPKDDAGWSPDHIAASAGRDEIVKALLINGAQVNAV
NONGCTALHYAASKNRHEIAVLMLEGGANPDANKHYDATAMHRAAAKGNLKMVHILIF
YKASTNIQDTGNTPLHLACDERVEEAKLLVTOGASIIYIENKEKTPLOVAKGGLGL
ILKRIAESEASM"

BASE COUNT 211 a 139 c 183 g 163 t
ORIGIN

Query Match 76.0%; Score 593; DB 94; Length 696;
Best Local Similarity 91.9%; Pred. No. 1.6e-169;
Matches 626; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

```
QY 97 atggagggggtgtgtctaaactaatgtctgcaacctggcctacagcgggaagctgaa 156
DB 1  ATGGAGGGGTGTGTCTAACCTAATGTGTCTTAACCTGGCCTACACGGGGAAGTGGAT 60
QY 157 agttgaaggagagtgattctggcgcataaaatccctggctactagaaactgaccagcagc 216
DB 61 GAGTTGAAGGAAGCATTTTGGCTGATTAAGTCTCTGGCCACTAGAACTGATCAGGACAGC 120
QY 217 aqaactgcattgcactgggcgtcagctgcagtgacacatacagaaattgttgaaattttgtg 276
DB 121 AGAACAGCATTCCTACTGGGCATGCTCAGCTGGTGCATACAGAAATTTGTAATTTCTTGTG 180
QY 277 caactggagtgccagtgaaatgataaaagacgatgcaggttggtctcctctcatattgcg 336
DB 181 CAACTTGGAGTCCCAATAATGAAAAGACGATGCGGTTGCTCTCCTTCATATTTGT 240
QY 337 gcttctgctggcgggatgagattgtaaaagccctctctggaaaagtgctcaagtgaat 396
DB 241 GCTTCGCTGGCGGGATGAGATTGTAAAGCCCTCTGTATAAAGGGGACACAGTGAAT 300
QY 397 gctgtcaatcaaaatggctgtactcccttacattatgcagcttcgaaaaaacagcgatgag 456
DB 301 GCGGTCAATCAGAAATGGCTGCGACGGCCCTTCATTATGACAGCTTCCAAAGAAATAGGATGAG 360
QY 457 atcgctgtcatattactggaagcgggctaatccagatgctaaaggccattatgagct 516
DB 361 ATTGCTGTTATGTTACTAGAGGTGGGGCTAATCCAGATGCTAAGNACCATTATGATGCT 420
QY 517 acagcaatcacccgggcagcagccaaagggttaacttgaagatgattcaatcctctctgtac 576
DB 421 ACAGCAATGCACCGGCAGCAGCAGCAAGGGTAACTTGAAGATGGTTTCATATCTCTTGTTC 480
QY 577 tacaagcatccacaacatccaaagacactgaggggttaaacctcctctacacttagcctt 636
DB 481 TACAAAGCATCCCAAAACATCCAAAGATGATGAGGGTAAACCTCTCTACACTTAGCTGT 540
QY 637 gatgaggagagtggaagaagcaaaactgctggtgtgtcccaaggagcaagattattacatt 696
```

```

|||||
Db 541 GATGAGGAGAGTGCAGAAAGCAAAATGCTGCTGACCAAGGACCAAGTATTACATT 600
|||||
Qy 697 gagaataaagaagaaagacacccctgcaagtggccaaagggtgacctggttgaataactc 756
|||||
Db 601 GAAATTAAGAGAGAAAGACACCGCTGCAAGTCGCCAAAGGGGGCCCTGGGTTTAATACTC 660
|||||
Qy 757 aagagaatggtggaaggttaa 777
|||||
Db 661 AAAAGAAATCGCAGAAAGTGAA 681
|||||

RESULT 5
AB022022
LOCUS AB022022 696 bp mRNA ROD 09-JAN-1999
DEFINITION Mus musculus mRNA for gankyrin, complete cds.
ACCESSION AB022022
VERSION AB022022.1 GI:4127259
KEYWORDS gankyrin.
SOURCE Mus musculus
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE Higashitsuji, H. and Fujita, J.
JOURNAL Submitted (03-JAN-1999) to the DDBJ/EMBL/GenBank databases. Hiroaki
Higashitsuji, Faculty of Medicine, Kyoto University, Clinical
Molecular Biology; Shogoin Kawahara-cho 54, Sakyo-ku, Kyoto, Kyoto
606-8507, Japan (E-mail: hhigashie@viru1.virus.kyoto-u.ac.jp,
Tel: 81-75-751-3753, Fax: 81-75-751-3750)
2 (bases 1 to 696)
REFERENCE
AUTHORS Higashitsuji, H. and Fujita, J.
TITLE Cloning of mouse gankyrin containing ankyrin repeats
JOURNAL Published Only in DataBase (1999) In press
FEATURES
source
1. .696
/organism="Mus musculus"
/db_xref="taxon:10090"
/dev_stage="adult"
/sex="female"
/tissue_type="placenta"
1. .696
/codon_start=1
/product="gankyrin"
/protein_id="BAA36969.1"
/db_xref="GI:4127260"
/translation="MEGCVSNIMICNLAYSGKLDKLERILADKSLATRTDQDSRTAL
HWACSGAHTIEVFELQLQVYPVNDKDGWSPLHTAASAGRDEIVKALLVKGHNSV
NONGCTPLHYAAKSRHETSVMLEGGANPDAKHDYDATAMHRAAKGNLKWHTLIF
YKASNIQDTGENTPLHLACDEERYEAKFLVTQGSIIYENKEEKTPLQVAKGGLGL
ILKRLAESEEAASM"
BASE COUNT 205 a 142 c 180 g 168 t 1 others
ORIGIN

Query Match 74.9%; Score 584; DB 94; Length 696;
Best Local Similarity 91.0%; Pred. No. 8.8e-167;
Matches 620; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 97 atggagggtgtgtctaacctaatggtctcaacctggctcaacctggctacagcggaagctggaa 156
|||||
Db 1 ATGGAGGGGTGTGTCTTACATATATGATCTGTAACTGGCCCTACAGTGGGAGCTGGAT 60
|||||

Qy 157 gagtgaagagagattgttggccgataaataccctggctactagaactgaccaggagcagc 216
|||||
Db 61 GAGTTCAAGGAGCGCATTTTGGCTGATAAATCTCTGCTACTAGAACTGATCAGGACAGC 120
|||||

Qy 217 agaactgattgcactgggcatgctcagctgacatacacagaattgttgaattttgttg 276
|||||
Db 121 AGAACAGCTTTGCACCTGGCGCATGCTCAGCTGGCCATACAGAAATTTGTAATTTCTGCTG 180
|||||

Qy 277 caacttggagtgccagtgatgaataaagacgatgcaggtgtgtctctctctcatattg 336
|||||

```

```

|||||
Db 181 CAACTTGGAGTGCCACTNAATGATAAAGATGACGCAAGTTGGTCTCTCTCATATTGCT 240
|||||
Qy 337 gcttctgtgctggccggatgagattgtaaaagccctctctgggaaaaaggtgctcaaggaat 396
|||||
Db 241 GCCTCGCTGGCGGGATGAGATTGTAAAGGCCCTTCTGTGTGAAGGTGCACATGTATAAT 300
|||||
Qy 397 gctgtcaatacaaaatggctgtactcctctacattatgcagcttcgaaaaacagggcatgag 456
|||||
Db 301 TCTGTCAATCAAAACGGCTGCACCTCCACTCCATTATGCAAGTTCGGAAGAAATAGGCATGAG 360
|||||
Qy 457 atcgctgtcatgttactggaagcgggggttaataccagatgctaaagaccatttatgagct 516
|||||
Db 361 ATTTCCTGTTATTGTTACTAGAAAGTGGGGCTAAACCCAGATGCGAAGCACCATACGATGCT 420
|||||
Qy 517 acagcaatcacccggcgagcagcccaagggttaactgaagatgattcatatcctctctgtac 576
|||||
Db 421 ACAGCAATCACCGGGCAGCAGCAAGGGTAACTTGAAGATGGTTCCACATCCCTTCTGTTC 480
|||||
Qy 577 tacaagcatcacaaaacatccaagacactgaggggttaacactctctacacttagcctgt 636
|||||
Db 481 TACAAAGCATCCACAAACATCCCAAGACACTGAGGTAACTCTCTACACTTAGGCTGT 540
|||||
Qy 637 gatgagagagagtggaagaagcaaaactgctggtgtcccaaggagcaagattattacatt 696
|||||
Db 541 GATGAAGAGAGAGTGGAAGAGGCAAAATTTCTGGTGACTCAAGGACCAAGTATTACATT 600
|||||
Qy 697 gagaataaagaagaaagacacccctgcaagtggccaaagggtgacctggttgaataactc 756
|||||
Db 601 GAGATAAAGAGAAAGAACACCCCTGCAAGTTGCCAAAGGGGGCCCTGGGTTTAATACTC 660
|||||
Qy 757 aagagaatggtggaaggttaa 777
|||||
Db 661 AAGAGACTAGCAGAAAGTGAA 681
|||||

RESULT 6
AC012149
LOCUS AC012149 180905 bp DNA HTG 06-SEP-2000
DEFINITION Homo sapiens chromosome 3 clone RP11-119E13, WORKING DRAFT
SEQUENCE, 12 unordered pieces.
ACCESSION AC012149
VERSION AC012149.8 GI:9438824
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Muzny, D.M., Adams, C., Bailey, M., Barberia, J., Blankenburg, K.,
Bodota, B., Bouck, J., Bowler, S., Brooks, A., Buhay, C., Bunac, C.,
Burkett, C., Burrows, J., Carter, M., Chacko, J., Chen, Z., Cox, C.,
David, R., Delgado, O., Deshazo, D., Ding, Y., Domah-Rashid, N.,
Dugan-Rocha, S., Durbin, K.J., Fernandez, C., Ferraguto, D.,
Forcum-Tansey, J., Frantz, P., Ganesh, R., Gorrell, J.H., Gorrell, L.L.,
Guevara, W., Harris, K., Hernandez, J., Hodgson, A., Hogues, M.,
Holloway, C., Hosak, H., Jackson, L.B., Jackson, L., Jia, Y., Jones, M.,
Kelly, S., Kondejewski, N., Kong, Y., Kovar, C., Leal, B., Li, Z.,
Lichter, O., Liu, J., Liu, W., Logan, O., Lozano, R.J., Lu, J.,
Lucier, R., Martin, R., Martinez, C., McLeod, M.P., Mei, G., Morgan, M.,
Morris, S., Nash, S., Nelson, A., Nguyen, R., Nguyen, N., Nguyen, S.,
Oswal, G., Parish, B., Paxton, S., Payton, B., Perez, L., Pu, L.L.,
Quiles, M., Reiter, D., Rives, M., Samuel, S., Say, J., Scherer, S.,
Shah, E., Shen, H., Simon, M., Sparks, A., Stamps, A., Sugeng, R.,
Tabor, P., Taylor, T., Vasquez, L., Vinson, R., Vo, Q., Wabnah, M.,
Watlington, S., Weinstock, G., Weinstock, I.R., Williamson, A.,
Worley, K., Wren, J., Wrensford, G., Yu, W., Zhou, X., Nelson, D. and
Gibbs, R.
DIRECT SUBMISSION
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 180905)
AUTHORS Worley, K.C.
TITLE Direct Submission

```


Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr20>
RP5-914P20 is from the library RPCI-5 constructed by the group of Pieter de Jong. For further details see

<http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2

This sequence is the entire insert of clone RP5-914P20 The true left end of clone RP5-955M13 is at 37950 in this sequence.

FEATURES

```

source
    Location/Qualifiers
        1..105334
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="20"
            /map="q13.13-13.2"
            /clone="rp5-914p20"
            /clone_lib="RPCI-5"
repeat_region
    1..212
        /note="AluJo repeat: matches 35..237 of consensus"
repeat_region
    682..985
        /note="AluX repeat: matches 1..304 of consensus"
repeat_region
    1184..1354
        /note="MIR repeat: matches 17..196 of consensus"
repeat_region
    1932..2066
        /note="FLAM_C repeat: matches 1..131 of consensus"
repeat_region
    2292..2398
        /note="L1ME3 repeat: matches 4915..5009 of consensus"
repeat_region
    2399..2692
        /note="AluJo repeat: matches 1..298 of consensus"
repeat_region
    2693..3111
        /note="L1ME3 repeat: matches 5009..5354 of consensus"
repeat_region
    3112..3407
        /note="AluX repeat: matches 1..297 of consensus"
repeat_region
    3408..3456
        /note="L1ME3 repeat: matches 5354..5396 of consensus"
repeat_region
    3506..3641
        /note="AluJb repeat: matches 163..294 of consensus"
repeat_region
    3662..3948
        /note="L1ME3 repeat: matches 5612..5929 of consensus"
repeat_region
    4569..4869
        /note="L1M4 repeat: matches 3660..3981 of consensus"
repeat_region
    5879..5998
        /note="MER5B repeat: matches 64..177 of consensus"
repeat_region
    6820..7120
        /note="AluSp repeat: matches 1..297 of consensus"
repeat_region
    7311..7457
        /note="AluJb repeat: matches 141..288 of consensus"
repeat_region
    8386..8684
        /note="AluSg repeat: matches 1..301 of consensus"
        complement(8685..9109)
misc_feature
    /note="match: GSS: Em:AQ151284"
repeat_region
    9525..9838
        /note="AluY repeat: matches 5..302 of consensus"
        complement(10462..10988)
misc_feature
    /note="match: GSS: Em:B89004"

```

```

misc_feature
/note="match: STS: Em:H18297"
misc_feature
complement(11204..11523)
/note="match: STS: Em:F06436"
misc_feature
complement(11209..11577)
/note="match: STS: Em:H79769"
repeat_region
11579..11614
/note="18 copies 2 mer aa 97% conserved"
mRNA
complement(join(11790..15959,23464..23556,25336..25448,
50098..50272,52273..52422))
/gene="dJ914P20.2"
/note="match: ESTs: Em:BE112660 Em:BE103176 Em:BF464091
Em:BF5853100 Em:AW139427 Em:BF445957 Em:AW173331
Em:AA111726 Em:BF318209 Em:AW389690 Em:BF523190"
/evidence=not_experimental
/product="dJ914P20.2.2 (activity-dependent neuroprotective
protein (K1AA0784, ADNP), isoform 2 (alternative 5' UTR))"
11790..12187
/note="match: STS: Em:G30243"
complement(11790)
/gene="dJ914P20.2"
complement(join(11790..15959,23464..23556,25336..25754))
/gene="match: cDNAs: Em:L25107 Em:AB020670 Em:AF068198
Em:AB018327
match: ESTs: Em:AI600029 Em:AA183241 Em:AA275189
Em:AA745620 Em:AA830325 Em:AI020367 Em:AA061696
Em:AI243490 Em:AI146218 Em:AV016003 Em:AA259672
Em:AI590352 Em:AA227896 Em:C14463 Em:244748 Em:AI071938
Em:AA874881 Em:AA495739 Em:D81665 Em:AI277858 Em:T34137
Em:AA371316 Em:AA067127 Em:R13866 Em:W32680 Em:AA067128
Em:W12764 Em:W45435 Em:AA528083 Em:AA373256 Em:AA249727
Em:AA790780 Em:243469 Em:AA636938 Em:C02493 Em:AI078094
Em:AI653383 Em:C15608 Em:AI605766 Em:AI636691 Em:R20042
Em:AA111726 Em:AV014718 Em:AA265030 Em:H26842 Em:AI548767
Em:AI394588 Em:R61133 Em:R23630 Em:AA243373 Em:AA739201
Em:AA207036 Em:AA600642 Em:AA243373 Em:AA739201
Em:AA291097 Em:AI007177 Em:T86563 Em:AI181625 Em:AI715802
Em:AA227596 Em:T84549 Em:AA874149 Em:AI1414826 Em:AA863656
Em:AI718275 Em:AI555864 Em:D44915 Em:AA354206 Em:AA722822
Em:AA495802 Em:AA806934 Em:W41625 Em:AI692233 Em:AI050875
Em:DI2066 Em:AA281988 Em:AI298135 Em:L44402 Em:AI655187
Em:AI674538 Em:AI548629 Em:AI548468 Em:T10165 Em:W40213"
/evidence=not_experimental
/product="dJ914P20.2.1 (activity-dependent neuroprotective
protein (K1AA0784, ADNP))"
complement(join(11790..15959,23464..23556,25336..25754,
50098..50272,52273..52422))
/gene="dJ914P20.2"
11791..12219
/note="match: STS: Em:T86741"
misc_feature
11796..12023
/note="match: STS: Em:T10164"
complement(11796)
/gene="dJ914P20.2"
complement(11814..11819)
/gene="dJ914P20.2"
complement(join(12852..15959,23464..23556,25336..25443))
/gene="dJ914P20.2"
/note="match: cDNAs: Em:AB018327"
/codon_start=1
/evidence=not_experimental
/product="dJ914P20.2.1 (activity-dependent neuroprotective
protein (K1AA0784, ADNP))"
/protein_id="CAB53748.2"
/db_xref="GI:13276226"
/db_xref="SPTREMBL:O94881"
/translation="MFOLPVNVLGSLRKARKTKVKILSDIGLEYCKEHIEDKPEPN
DFYLNKNTWEDVGLWDPSLTKNQDYTKPCCSCAPFSKPFSAVKSHFNVKSDFE
NRILLNCPYCFNADLKYLETHIKIFHAPNAPSSSLSTFDKXKNDGLPKQADSV
EQAVYKCKTFRDPLYEIVRKHYIHFQHVAAPIYIAKAGEKSLNGLPKQNSV
SSIHCKRCLFMPKQVHEDHRIQVYGTAMIGHTNVVVRSPKPLIAPKPQ
DKKSMGLPPIRIGSLASGNVRSLSQOQVNRSLIPKPNLNSTGVNMSSVHLQNNYGV

```



```
KSVGGYXVSGOSMRGLGPNAPVSTPOOSVSKVOLLPSNGRSYGLGSEORSOAPARY
SLOSANASLSGQKSPSLQSOSQASRVLGSSSKPAAATGPPGNTSSSTQKWKICT
ICNELFPENVSFHEPEKHEKAEKPAVANYIMKIHNTSKCLYCNRLPTDNLNHL
IHGLSCPYSTFNDVEKMAHMRVHIIDEEMPKTSTLSFLDTLQOGSHTNLHLV
TNYNLRDAPAESVYHAQNRPVPKQPKQKADIPKSPQAAVYKPKDVGKTLIC
PLCFSLKGPISDALHLRHRHVIQVHPVKRLTYKICHLGCVVTSNMTASTITL
HIVHCRGVKQNGODKTNAPSLRNLSPSLAPVKRTYBOMEFFPLKKRKLDDSDSPS
FEEKPEPEVVLALDPKDHEDDSYARKSLFKTKYFNKOPYPTRELEKLAASLWKS
DIASHFSNKRKVCVRCKYKPGVLLGFNMKLNKVKHEMDFDAEWLFENHDEKDSRV
NASKTADKLNGLKEDDSSDSFENLEESNESGSPFPDFVEFEPKISNDNPEKGLS
VIPEDASEEKLKDGDSKYETIHLTEPTKLMHNASDSVDDDDVVEWKGASLP
ESGPGSQVSPEDNTECMKPGTWSDESSDARSSKPAAKKATGMDGDEQLKWN
SSYGVGEFWSKDSQKSNWNASENDERLSNPQIEWQNSTIDSEDEGEQFQDNMTDGVAEPM
HSLAGVKLSSQQA"
misc_feature
complement(13731..14156)
/gene="d0914P20.2"
/note="match: STS: Em:G37216"
13814..14173
misc_feature
/note="match: STS: Em:L44402"
16060..16350
repeat_region
/note="AluSq repeat: matches 1..291 of consensus"
16607..17008
repeat_region
/note="Tigger3(Golem) repeat: matches 2649..3028 of consensus"
17009..17274
repeat_region
/note="AluJo repeat: matches 1..286 of consensus"
17275..17389
repeat_region
/note="Tigger3(Golem) repeat: matches 2542..2649 of consensus"
Query Match 68.2%; Score 532; DB 92; Length 105334;
Best Local Similarity 89.9%; Pred. No.1.5e-150;
Matches 629; Conservative 0; Mismatches 60; Indels 11; Gaps 5;
Qy 82 tgcctggagacagcaaaatggaggggtgtgtcttaacctaatgtctgcgaacctggcctac 141
Db 35160 TCCTGGGAAAGCAACAGACGGGTGTCTCTAAACTTTATGTCTGTCAACTTTGGCCTAC 35219
Qy 142 agcgggaagctggaagagtgatctctgcccataataatccctggctactaga 201
Db 35220 AGAGGGAAGCTGGAGAGGTGAAGGAGAGATCTGGCCATAAACCCTGGCTATGAGA 35279
Qy 202 actgaccagacagcaaatgcattgcactgggcatgctcagctggacatacagaatt 261
Db 35280 ACTGACCAGGA---CAGAAGTGCATTGCATTGGCATGCTCAGCCATACATACATAAAT 35336
Qy 262 gttgaattttgtgcaactggagtgccagtggaatgataaagaacgatgc-aggttgctc 320
Db 35337 ATTGAATTTTGTTCACCTTGGAGTGCCAGTGAATGATAAAGACTATGCGAGATTGGTC 35396
Qy 321 tcctcttcattatggctctctgctgcccggatgagatgtgaaagccctcttggaata 380
Db 35397 TCCTCTTCAGACTGCTGTTGTGTCGCCAGATGACACTGTAAGCCCTCTGGGAAA 35456
Qy 381 aggtgctcaagtgaatgctgcaataaaatggctgtactcccttaccattatgcagcttc 440
Db 35457 AGGTGCTCAAGTGAATGCTGCAATCAAAATGGCTCTACTCCCTACATCATGAGCTTC 35516
Qy 441 gaaaaaaggagatgagatcgctgtcttactggaaggcgggctaatccagatgctaa 500
Db 35517 CAAAACAGGATGAGATGCTGCTCATGTTACTAGAGAGGTGGGCTAATCCAGATGTAA 35576
Qy 501 ggaccattatgagctacagcaatgacccggcagccagcaagggttaacttgaagatgat 560
Db 35577 GGATCATTTATGAGGCTAGCAAGACAGCCAGCCACCAAGGGTAACTTCATCAAGATGAT 35636
Qy 561 tcatctctctgtactacaagcatccacaacatccaaagacactgagggtaacatcc 620
Db 35637 TCATATCTCTGTACTACAAGCATCCACAAATCATCATCAGACACTGAGGGTAACTCC 35696
Qy 621 tctaacacttagccttgatgagagagagtggaagagcaaaaactgctggtcccaagg 680
Db 35697 TCCACACTTAGTCTGTGAT-----AAGAGTGGAAAGAGCAAAACTGCTGTGTCCTCCCAAGG 35751
```

```
Qy 681 agcaagtatttaccattagaataaagaagaacacacccctgcaagtggccaaagtgg 740
Db 35752 AGCAA-TATTTACATAGATAAAGAAGAAAGAACCCCA-ACAACTGGCCAAAGGTGC 35809
Qy 741 cctgggttatactacagaagaatggggaagggttaaca 780
Db 35810 CCTGGGT-TAGTACTCAAGAGAATGGTAGAAGGTTAAACA 35849
RESULT 8
AC009863/c
LOCUS AC009863 175550 bp DNA HTG 26-MAY-2000
DEFINITION Homo sapiens chromosome 3 clone RP11-573D15 map 3, WORKING DRAFT
SEQUENCE, 16 unordered pieces.
AC009863
AC009863.3 GI:8096876
HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 175550)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
JOURNAL Homo sapiens chromosome 3, clone RP11-573D15
TITLE Unpublished
AUTHORS 2 (bases 1 to 175550)
REFERENCE 1 (bases 1 to 175550)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Depayre,E., Devon,K., Dewar,K.,
Donelan,L., Doyle,M., Ferreira,P., Fitzhugh,W., Forrest,C.,
Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
Karatas,A., Lehoczy,J., Lieu,C., Locke,K., Macdonald,P.,
Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
Meldrum,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
Testaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
Direct Submission
Submitted (03-SEP-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 26, 2000 this sequence version replaced gi:6514025.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L1835
Center clone name: 573_D_15
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 139143 bases at least Q40
Consensus quality: 158720 bases at least Q30
Consensus quality: 168324 bases at least Q20
Insert size: 182000; agarose-fp
Insert size: 174050; sum-of-contigs
Quality coverage: 4.0 in Q20 bases; agarose-fp
Quality coverage: 4.1 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
```


During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Embl, EMBL; Swi, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/projects/C.elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome 13, constructed by the Sanger Centre Chromosome 13 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr13>

RP11-505F3 is from the library RPCI-11.2 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/> VECTOR: pBACE3.6

This sequence is the entire insert of clone RP11-505F3.

FEATURES
source
1. 199882
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
/clone="RP11-505F3"
/clone_lib="RPCI-11.2"
/clone_1ib="RPCI-11.2"

BASE COUNT 58871 a 40074 c 40071 g 60866 t

ORIGIN

Query Match 51.1%; Score 398.2; DB 90; Length 199882;
Best Local Similarity 93.6%; Pred. No. 9.8e-110;
Matches 437; Conservative 0; Mismatches 28; Indels 2; Gaps 2;

QY 79 agttgtggagcagaaatggaggggtgtgtgtacctaattgtgtcgaacctggcc 138
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 101505 AGTTCTCTGGGAAGCAAAATGGAGGGGTGTGTCTTAACCTAATGCTGCAATCTGGCC 101564

QY 139 tacagcgggaagcgtgaagaggttgaaggagagattcttgcgcgataaatccctggctact 198
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 101565 TACAGCAGGAAGCTGGAAGAGTTCAAGGAGGGTATTCTGCGCGATAAATCCCTAGCTACT 101624

QY 199 agaactgaccagcagcagcaactgcattgcactggcgatctcagctgggacacacagaa 258
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 101625 AGAAGTACACGAGCAGCAGACAGCATTGCTAGGCGATGCTCAGCTGGACATAAGGA 101684

QY 259 attgttgaaattttgttgcactggagtgccagtgaatgaataaagacgatcgaggttgg 318
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 101685 ATTGATGAATTTTGTGCAACTTGGAGTGGCAGTGAATGATAAAGACGATGCAAGTTGG 101744

QY 319 tctctcttcattatcgccgtctctgcgcggagatgagattgataaagccctctggga 378
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 101745 TCTCTCTC-TCAAAATGCGCGCTTCTGCTGCGCGGATGAGATTGATAAAGCCCTCTTGGA 101803

QY 379 aaagtgctcgaagtgaatgctgtcgaatcaaaatggctgtactcccttaccattatgcagct 438
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 101804 AAAGTGCTCAAGTCAATGCTGCTCAATCAAAATGACTACTCTCCCTACATTATGCAGCT 101863

QY 439 tcgaataaacagcagcagatgagatcgctgtcattgtactggaagcggcgctaaatccagatgct 498
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 101864 TCCAAAATAGGCATGAGATTGCTGTCTGTTACTAGAGGCGGGCTTAATCCTCATGCT 101923

QY 499 aaggaccattatgagcgtacagcaatgcaccggcgagcagcagcaagg 545
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 101924 AAGGA-CATTATGAGGCTACAGCATTGCTAGGAGGAGGCCAAGGG 101969

RESULT 10
AC072018
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AC072018 145483 bp DNA HTG 06-SEP-2000
Homo sapiens chromosome 3 clone RP11-24009, *** SEQUENCING IN
PROGRESS ***, 40 unordered pieces.
AC072018
HTG: HTGS_PHASE1.
AC072018.3 GI:9966538
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 145483)
Muzny,D.M., Adams,C., Bailey,M., Barbara,J., Blankenburg,K.,
Bodota,B., Bouck,J.J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hoques,M.,
Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,
Kelly,S., Kondejewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z.,
Lichtarge,O., Liu,J., Liu,W., Logan,O., Lozado,R.J., Lu,J.,
Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M.,
Morris,S., Nash,S., Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S.,
Oswal,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L.,
Quiles,M., Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,S.,
Shah,E., Shen,H., Simon,M., Sparks,A., Stamps,A., Sucgang,R.,
Taber,P., Taylor,T., Vasquez,L., Vinson,R., Vo,Q., Wahbah,M.,
Watlington,S., Weinstein,G., Weinstein,I.R., Williamson,A.,
Worley,K., Wren,J., Wrensford,G., Yu,W., Zhou,X., Nelson,D. and
Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 145483)
Worley,K.C.
Direct Submission
Submitted (07-JUN-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Sep 5, 2000 this sequence version replaced gi:8572872.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HAHU
Center clone name: RP11-24009
----- Summary Statistics
Sequencing vector: M13; L08821
Chemistry: Dye-terminator Big Dye; 93% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 112126 bases at least Q40
Consensus quality: 127972 bases at least Q30
Consensus quality: 136731 bases at least Q20
Estimated insert size: 134186; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 2.4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 40 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 10264: contig of 10264 bp in length

* 10265 10364: gap of unknown length
* 10365 17786: contig of 7422 bp in length
* 17787 17886: gap of unknown length
* 17887 24658: contig of 6772 bp in length
* 24659 24758: gap of unknown length
* 24759 30985: contig of 6227 bp in length
* 30986 31085: gap of unknown length
* 31086 36928: contig of 5843 bp in length
* 36929 37028: gap of unknown length
* 37029 43546: contig of 6518 bp in length
* 43547 43646: gap of unknown length
* 43647 50051: contig of 6405 bp in length
* 50052 50151: gap of unknown length
* 50152 55816: contig of 5665 bp in length
* 55817 55916: gap of unknown length
* 55917 62446: contig of 6530 bp in length
* 62447 62546: gap of unknown length
* 62547 67354: contig of 4808 bp in length
* 67355 67454: gap of unknown length
* 67455 72321: contig of 4867 bp in length
* 72322 72421: gap of unknown length
* 72422 76988: contig of 4567 bp in length
* 76989 77088: gap of unknown length
* 77089 80703: contig of 3615 bp in length
* 80704 80803: gap of unknown length
* 80804 84396: contig of 3593 bp in length
* 84397 84496: gap of unknown length
* 84497 88793: contig of 4297 bp in length
* 88794 88893: gap of unknown length
* 88894 92423: contig of 3530 bp in length
* 92424 92523: gap of unknown length
* 92524 95759: contig of 3236 bp in length
* 95760 95859: gap of unknown length
* 95860 99818: contig of 3959 bp in length
* 99819 99918: gap of unknown length
* 99919 103132: contig of 3214 bp in length
* 103133 103232: gap of unknown length
* 103233 105760: contig of 2528 bp in length
* 105761 105860: gap of unknown length
* 105861 108187: contig of 2327 bp in length
* 108188 108287: gap of unknown length
* 108288 111981: contig of 3694 bp in length
* 111982 112081: gap of unknown length
* 112082 114420: contig of 2339 bp in length
* 114421 114520: gap of unknown length
* 114521 116992: contig of 2472 bp in length
* 116993 117093: contig of unknown length
* 117093 119463: contig of 2371 bp in length
* 119464 119563: gap of unknown length
* 119564 121783: contig of 2220 bp in length
* 121784 121883: gap of unknown length
* 121884 123681: contig of 1798 bp in length
* 123682 123781: gap of unknown length
* 123782 126198: contig of 2417 bp in length
* 126199 126298: gap of unknown length
* 126299 128556: contig of 2258 bp in length
* 128557 128656: gap of unknown length
* 128657 129817: contig of 1161 bp in length
* 129818 129917: gap of unknown length
* 129918 131742: contig of 1825 bp in length
* 131743 131842: gap of unknown length
* 131843 133420: contig of 1578 bp in length
* 133421 133520: gap of unknown length
* 133521 134903: contig of 1383 bp in length
* 134904 135003: gap of unknown length
* 135004 136266: contig of 1263 bp in length
* 136267 136366: gap of unknown length
* 136367 137771: contig of 1405 bp in length
* 137772 137871: gap of unknown length
* 137872 138996: contig of 1125 bp in length
* 138997 139097: gap of unknown length
* 139097 140945: contig of 1849 bp in length
* 140946 141045: gap of unknown length

* 141046 142678: contig of 1633 bp in length
* 142679 142778: gap of unknown length
* 142779 144032: contig of 1254 bp in length
* 144033 144132: gap of unknown length
* 144133 145483: contig of 1351 bp in length.
FEATURES
Location/Qualifiers
1..145483
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-24009"
/c31243 g 31243 g 40572 t 4037 others
BASE COUNT 38303 a 31328 c 31243 g 40572 t 4037 others
ORIGIN
Query Match 28.7%; Score 223.8; DB 74; Length 145483;
Best Local Similarity 91.9%; Pred. No. 1.3e-56;
Matches 249; Conservative 0; Mismatches 17; Indels 5; Gaps 1;
QY 510 tgaggtacagcaatgcaccggcgagcagccagggtaactgaagatgatcatatcct 569
|||||
Db 119572 TGAGGCTACAGCAATGCACCGGTACAGCAGCCAGGGTAACCTTGAGATGATTCCTATCCT 119631
|||||
QY 570 tctgtactacaagaatccacaacatccaaacacactgagggtaaacactcctctacact 629
|||||
Db 119632 TCTATACTACAGCAATCCACGAACATCCAGACACTGAGGGTAACACTCCTCTACACTT 119691
|||||
QY 630 agcctgtgatgaggagagagtggagaagcaaaaactgctggtgctcccaaggagcaagtat 689
|||||
Db 119692 AGCCTGTGATGAGGAGAGAGTGGAAAGGA-----GCAAGTGTCCCAAGGAGCAAGTAT 119746
|||||
QY 690 ttacattgagaataagaagaacacacccctgcaagtgcacaaagtgctcctgggttt 749
|||||
Db 119747 TCACATGGAGAATAAAGAAAGAAAGACACCCCTGCAGCGCCAGAGGTGGCTACGTTT 119806
|||||
QY 750 aatactcaagaagaatggtagaggttaaa 780
|||||
Db 119807 AATACTCAAGTCAATGGTGGAGGTTAAACA 119837
|||||
RESULT 11
HS889N15/c
LOCUS HS889N15 123395 bp DNA PRI 23-NOV-1999
DEFINITION Human DNA sequence from clone 889N15 on chromosome Xq22.1-22.3.
Contains part of the gene for a novel protein similar to X.laevis
Cortical Thymocyte Marker Ctx, the possibly alternatively spliced
gene for 26S Proteasome subunit p28 (Ankyrin repeat protein), a
novel gene and exons 36 through 45 of the COL4A6 for Collagen Alpha
6(IV). Contains ESTs, STSS, GSSs and a putative CpG island,
complete sequence.
ACCESSION AL031177
VERSION AL031177.1 GI:4071056
KEYWORDS HTG; 26S Proteasome subunit p28; Ankyrin repeat; COL4A6; Collagen
Alpha 6(IV); Cortical Thymocyte Marker; CpG island; Ctx.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 123395)
AUTHORS Grafham, D.
TITLE Direct Submission
JOURNAL Submitted (06-JAN-1999) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
COMMENT requests: clonerequest@sanger.ac.uk
On Dec 29, 1998 this sequence version replaced gi:4034456.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence is the entire insert of clone 889N15. This sequence
has been finished according to sequence map criteria as follows. An

RESULT	13											
DROANKY		DROANKY	5069 bp	DNA	INV	25-OCT-1994						
LOCUS		Drosophila melanogaster ankyrin mRNA, complete cds.										
DEFINITION		L35601										
ACCESSION		L35601.1	GI:557083									
VERSION		ankyrin.										
KEYWORDS		Drosophila melanogaster.										
SOURCE		Drosophila melanogaster										
ORGANISM		Drosophila melanogaster										
		Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;										
		Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;										
		Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;										
		Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.										
		1 (bases 1 to 5069)										
REFERENCE		Dubreuil,R.R. and Yu,J.										
AUTHORS		Ankyrin and beta-spectrin accumulate independently of										
TITLE		alpha-spectrin in Drosophila										
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 91 (22), 10285-10289 (1994)										
MEDLINE		95024098										
FEATURES		Location/Qualifiers										
		1..5069										
		/organism="Drosophila melanogaster"										
		/isolate="All"										
		/isolate="A631"										
		/db_xref="taxon:7227"										
		/clone_lib="Nick Brown and Carl Thummel"										
		/clone="Nick Brown and Carl Thummel"										
		1..5069										
		1..33										
		34..4683										
		/codon_start=1										
		/product="ankyrin"										
		/protein_id="AAC37208.1"										
		/db_xref="GI:557084"										
		/translation="*MTLGDVVRVENKIQTSETCAINGMALDNKNGIIKONDATISFLR										
		AKSGDIKVMFDLDCGEISDINSCNANGNLNHLIAAKDGYVDICELRRGKIKDNA										
		TKKGTNLTIASLAGOHVDINILYNANVNVQSLNGFTPLVMAAGAHNDHNCRTLLA										
		NGANPSLSTEDGFTPLAVAMQGHDKIVAVLLENDVRGKVRPLPALHIAAKNDVNAAK										
		LILLOHPNADIVSKSGFTPLHIAAHYGNVDIATLLLNKADVNYVAKHNTPLHVAKL										
		WKLISCTLLLCAGAKIDAATRDGLTPLHCASRSHGHEVIKHLLOQNAFILTKNGL										
		SALHMAAQGEHEAAHLHNKAPDEVVDYLTALHVAAGHGVKVALLDYIKRANP										
		NARALNGFTPLHIAACKNRIKVELLIKHGANGITATESGLTPLHVASFMGCIINIVY										
		LHQEASADLPTIRGTPLHIAARANQADIIIRLSAKVDAIVREGQVTPLVHVASLG										
		NTNIIMLLQHGAEINQAQSDNYSALHIAAKEGOENIYOVLLENGEANNATVKKSGFT										
		LHLACKYKONVVOILLONGASTDFQGNKDVTPLVATHYNNPSIVELLELLKNGSSPNL										
		CARNGQCAIHACKKNYLEINQQLQHGADVNIISKSGESPLHIAAQGNGVDVQLL										
		EYGVISAQAQKTPLVHAAQGLVDVLSQLLEHGANIISERTNGTTPLVHMAHYGHL										
		DIKVFTEINDADTMSNSTGYTPLHQAQOQHIMIINLLHRKANPALTUKDGTALH										
		IASNLGYVTWMSLKIVTSTVINSNIGAEIEKLVKWTPELMOETLLSDSDSCDDL										
		LQHNHYVMATDVLKANGYODOKNFTTNDHLDTVSNLKNKEILLPNMCSIELTEIPI										
		GKPDNVNARSOVHLGFVSLVDARGGSMRGYRNGVRIIVPPKCAEPRTICRY										
		VKQVRVNPPLMEGALVSLILEMSPVDGMFLSPITLVEPHYGTILKNEREILLIRS										
		DNGSREHNLXYKDIIEDINQTEEFSDRIVRIVTQNPHFPFVAVSRVQEVHVHIGP										
		GDGTVESTAVPQKILFPPHATKIRIVGLQASVDLVECSKLLGGQVAVSPVTVTG										
		RRKFLFKATLSTIPAKACTNSVMNACYGNSSPSTLRLLCSISGGQRAFWEDVYTG										
		STPLSFVRDSTVFTTTSARFWLIDCRNII DAGRMATELYSHLAKVPFYVKVFIKAR										
		ISQTEAKSFVCTMDKEDKTLQOQSYFTEVAKSRDIEVLQNGIIVLEFAGNVLPIIK										
		GKQLTIFQPCENLRSFAHKDQEFPHGRICFTMYPMWGDEVPPLPCLINLSY										
		DKFTITNHLRDLNUNICINAHGKLNHNENIVFGVKQOVKKKIDITKACINMSDDL										
		LIRHEADVILDDICSHLGSOWPLLVGLVQSDIDLVKTFLEFLNDSVKOSMAMQLML										
		EHGGILTGNVLAELYKIGRSIVDEKSFNAEFGTHQPEKVLPTAGTIGKGDGYEQ"										
		4684..5069										
3'UTR		1612 a	996 c	1048 g	1413 t							
BASE COUNT												
ORIGIN												
Query Match		14.8%;	Score 115.6;	DB 6;	Length 5069;							
Best Local Similarity		51.1%;	pred. No. 6.8e-24;									
Matches 271;		Conservative	0;	Mismatches	259;	Indels	0;	Gaps	0;			
Qy	202	actgaccaggacagcaactgcatctggcgatctgcagctgcagacatacagaatt	261									
Db	1510	ATTCTCGTGAAGCCAAACGCCACTTCATGTAGTCTCTCATTTAGGAATAATTAATA	1569									

```
QY 262 gttgaattttttgttgaacttgagtgccagtggaatgataaaagacgatgcaggttggtct 321
    || || || || || || || || || || || || || || || || || || || || ||
Db 3828 ATTATGTTATTACTGCAGCATGCAGCAGAGATAAATGCTCAATCTAATGACAAATATTGC 3887
QY 322 cctcttcataatgctgctctctgctgcccgggagtgagattgtaaaagccctcttggaataa 381
    || || || || || || || || || || || || || || || || || || || || ||
Db 3888 GCGCTGCATATTGCTGCAAAAGAAGCACAAGAGATAATAGTCCCAAGTGTACTGGAAAT 3947
QY 382 ggtgctccaagtgaatgctgctcaatcaaatgctgtactctcccttcacattatgcagcttcg 441
    || || || || || || || || || || || || || || || || || || || || ||
Db 3948 GTGCGGAAATATAGCTGTTACCAAAAAGGGTTTACTCCACTGCTGATTTACTTGCARA 4007
QY 442 aaaaacagcgatgagatgcgtgctatgcttactggaagcggggctcaatccagatgcgtaag 501
    || || || || || || || || || || || || || || || || || || || || ||
Db 4008 TAGCGAAACGACAGATGTTGCTTCAATATATTACTACAAATGGAGCTTCTATTGATTTTCAG 4067
QY 502 gaccattatgagctacagcaatgacccgggagcagcagccaaagggtaacctggaagtatt 561
    || || || || || || || || || || || || || || || || || || || || ||
Db 4068 GGGAAAAATGATGTGACCCCATTTACACGTGGCCACGCACATACAATAATCCCTCAATAGTA 4127
QY 562 catactctctgtactacaagaatccacacacacacacacacacacacacacacacactct 621
    || || || || || || || || || || || || || || || || || || || || ||
Db 4128 GAACCTCTTTTAAAAAATGGTGCATCGCCCAATTTATGTGCCGAATGGTCAATGTGCT 4187
QY 622 ctacacttagctgtgatgagagagagtggaagaagcaaaactgctgtgtgtcccaagga 681
    || || || || || || || || || || || || || || || || || || || || ||
Db 4188 ATTACATAGCTTGCAAAAAAATATCTGGAATTTGCGATGCAGCTGTTGCAACACGCT 4247
QY 682 gcaagtattacattgagataaagaagaagaaagacacccctgcaagtggc 731
    || || || || || || || || || || || || || || || || || || || || ||
Db 4248 GCAGACGTTAATATTATTAGCAAGAGTGGCTTCTCTCCGTTACATTTGCG 4297
```

RESULT 15

```
AE003845/c
LOCUS
DEFINITION
  Drosophila melanogaster genomic scaffold 142000013385266 section 1
  of 3, complete sequence.
ACCESSION
  AE003845 AE002796
VERSION
  AE003845.2 GI:10726333
KEYWORDS
  HTG.
SOURCE
  fruit fly.
  Drosophila melanogaster
  Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
  Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
  Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
  Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
  Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F.,
  George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,
  Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X.,
  Brandon,R.C., Rogers,Y.H., Blazer,R.G., Champen,M., Pfeiffer,B.D.,
  Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor
  Miklos,G.L., Abril,J.F., Agbayani,A., An,H.J.,
  Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Basu,A.,
  Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y.,
  Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borkova,D.,
  Botchan,M.R., Bouck,J.J., Brokstein,P., Brottier,P., Burtis,K.C.,
  Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I.,
  Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de
  Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M.,
  Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C.,
  Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferriera,S.,
  Fleischmann,W., Foster,C., Gabrielian,A.E., Garg,N.S.,
  Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,
  Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J.,
  Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J.,
  Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z.,
  Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,
  Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A.,
  Li,J., Li,Z., Llang,Y., Lin,X., Liu,X., Mattel,B., McIntosh,T.C.,
  McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C.,
```

```

Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L.,
Muzny,D.M., Nelson,D.L., Nelson,D.R., Palazzolo,M., Pittman,K.,
Nusskern,D.R., Pacle,J.M., Palazzolo,M., Pittman,K., Pan,S.,
Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K.,
Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I.,
Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C.,
Stapleton,M., Strong,R., Sun,E., Svirskaas,R., Tector,C., Turner,R.,
Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A.,
Weinstock,G.M., Weissbach,J., Williams,S.M., Woodage,T.,
Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F.,
Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H.,
Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,
Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.
The genome sequence of Drosophila melanogaster
Science 287 (5461), 2185-2195 (2000)
20196006
2 (bases 1 to 114334)
Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.
Direct Submission
Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
On Oct 9, 2000 this sequence version replaced gi:7304337.
COMMENT
  Location/Qualifiers
  FEATURES
    source
      1. .114334
        /organism="Drosophila melanogaster"
        /db_xref="taxon:7227"
        /chromosome="4"
        complement(join(<1. .1937,3565. >4357))
        /gene="plexB"
        /note="Nucleotide sequence of the Celera sequence differs
        from the published sequence for this transcript."
        /product="CT38225"
        /db_xref="FLYBASE:FBan0017245"
        /db_xref="FLYBASE:FBgn0025740"
        complement(<1. >4357)
        /gene="plexB"
        /note="CG17245"
        /map="101F1-102A1"
        /db_xref="FLYBASE:FBan0017245"
        /db_xref="FLYBASE:FBgn0025740"
        complement(join(<1. .1937,3565. .4075))
        /gene="plexB"
        /note="plexB gene product; Nucleotide sequence of the
        Celera sequence differs from the published sequence for
        this transcript"
        /codon_start=1
        /db_xref="FLYBASE:FBan0017245"
        /db_xref="FLYBASE:FBgn0025740"
        /protein_id="AAF59374.1"
        /db_xref="GI:7304343"
        /translation="MLRKELVCANIHLVHLVLCIIILGSHQYCNVCELPAAALSP
        LNDIVAQFNLPSPVQSSPASSYGNRSIGNIESVRDSQKNYFTHMSFDFMHNVLFAQ
        ATNKLKLNENLRVLAENAVTGPLHDSQCHAGCPEDTETSLVNNFNKLLVSVYADG
        ILIACGSIROGACEIYSLPRFPATPOFFAVPLAANDENASTYFVGPARYAKKEDIL
        YSGYFTFNVDGRHDPVPAISSRDLDDLNIAEFSIQSDINIDVKYRDHLDYDYGFN
        RDVCTPASHKLAOKMKGKDDHVLVTVPSPRSNQPESKAMCIYSIKDIEPMFI
        ENHLCNGTKDKRNLGVISCTINDGRCPIVSGSLGNIYFCSVGLKISGVSPITAHIL
        FHFNVSVTSTATSTDOOHSFLAFGTNMGVIRKVLGSGSPGEYEIVVDAGNRIL
        PTQMSPKKDLVLSQRKITKRIEHCSTVNTCSACLESRDPPFGWCSLEKRCRTVRS
        NCSDTSASRWLSLGSQQCIETESIPERIKIPELSHLHLIIRTPPPFNKRVCF
        GNSYPIDAEILGGLCATPDLRPLTNTDHLVPLSVRSSTKDFVSRNFAFF
        CDSHHGNCQECLOQSSWGCNWCIFDNKCVKHSLOCRNIENAVSTVGHCPHLKKNRPEIL
        LPVRVPIEIRLEIENLPKPSKHAGFLCTIHEAAQMLLPAHVSNKIVCEKTPYFY
        EINTHEQAKVVVTWNQF"
        complement(join(8359. .11113,11167. .11331,11396. .11908,
        .11965. .12605,12668. .13012,16484. .16917))
        /gene="ci"
        /note="Nucleotide sequence of the Celera sequence differs
        from the published sequence for this transcript."
        /product="CT6841"
        /db_xref="FLYBASE:FBan0002125"
        /db_xref="FLYBASE:FBgn0004859"
```



```
gene
complement(<8359. >16917)
/feature="ci"
/feature="CG2125"
/feature="101F1-102A1"
/feature="FLYBASE:FBan0002125"
/feature="FLYBASE:FBgn0004859"
complement(join(8645. .11113,11167. .11331,11396. .11908,
11965. .12605,12668. .13012,16484. .16526))
/feature="ci"
/feature="ci" gene product; Nucleotide sequence of the Celera
sequence differs from the published sequence for this
transcript
/codon_start=1
/feature="FLYBASE:FBan0002125"
/feature="FLYBASE:FBgn0004859"
/protein_id="AAF59373.1"
/feature="GI:7304342"
/translation="MKWTPRYLHIFLPSRRAAAVAAATVLPSPCINQHPTDVS
SSVTPVPIPTGTSISIKTSIQPOICNENTLILGNAGHOHNHQPQHVHINVTGPHD
FHPAYRIPGMQLYSLQNTSASSPHDPPVNCASAFHLAGLGLGSADFLGSRGLSSL
GELHNAVAAAGASLASLDFHESVDGNERLGSPPGSGIRASLSKRALSSPSYSD
SFDSINMLRPSNLSLAINNGSGSAASGSGHISATVAINPMHSHVSTRLOQIAHL
LRASAGLIMPPTQQAASGFSIGHMPTASLRVNDVHNLSDSHIQTITSPVTKDV
SOVPAFAFLNLDLAREKKGPKVDVPEQPSSTGGVAQVEADSSOLSQRXNNV
VNNITGPGDKVNSRLDEYINGSGISIPSEYDCANATTDIKDEPGDFIETNCHWR
SCRIETIDELVKHINNDHIOTNKFAVCRWEDCTRGKPFKAQVLMVHMRHTGE
KPHKTFEGCFKAYSLENLKTHLRSHTGKPYTCYPCGSKAFNASDRKHQNRTH
SNEKPYICAPGCTKYTPDPSLRKHVTVHGAEFYANKKHKGLPLINDRANSRLQONS
RHNLDHNDISSPSCSDSLHGLMGLHSSPISKESDISSNHLHLYNGVRSASLITYS
PDALAENLDNWDNDVDDVADLPIVLRAMVINGNASASTIGSVLAKORFGR
LOTKGNSTILMCLNIPESNRTFGISELNRITELMEPTGDAETIKPLNPTITGGY
TEDPLQNTSFRNTVSKQSTGSGTQGOFRDQNSTASTYYSQMSRRSSOSQSV
SIPTRPNPSCNSTAFEDYIPSPGCSRSSOMNGANCSFTSTGLPLVKNESKSL
NACINKPNIGVOGVLYNSLPPSPSSHLIATNLKLRKDSFYHNFSTGRFSVPSY
HSIHIKNNPVGNEDEKAIASNARQTDVFNINLDPLTNSRSTTTPHSDINWGK
TNNIASINKNDLRKDLFTVSIKADAMTSDQHPNERINLDEVEELILPEMLQYLNL
VKDTHLNDKQVGVNSVSETIASNHYREGSNITYNKQILLPPSNVDIOPNTTK
FTVQDFAMTAVGGSFQRELSLTPVNEHGHAKCESFHQSQKYNMTDIGKSQSL
PSAHOKEKSNYNOIIDSMTSLPELNVDSIFRNENIFKVDHNDHNEIQCGIS
QSQSPSTNLNDQSFSTVNMOPITTSKLPPEPOKIVGDTQASNTSVMLHDTYQTL
EYVQSCONWETNTSNQIQSLPGMPVNTLPFDVSSSTHPYHGTNMVINDMTTSLT
SLLEENRYLQMMQ"
complement(join(26773. .27392,27494. .27890))
/feature="Rps3A"
/feature="Nucleotide sequence of the Celera sequence differs
from the published sequence for this transcript."
/product="CT7064"
/feature="FLYBASE:FBan0002168"
/feature="FLYBASE:FBgn0017545"
complement(<26773. >27890)
/feature="Rps3A"
/feature="CG2168"
/feature="102A1-102A1"
/feature="FLYBASE:FBan0002168"
/feature="FLYBASE:FBgn0017545"
complement(join(26927. .27392,27494. .27834))
/feature="Rps3A"
/feature="Rps3A" gene product; Nucleotide sequence of the
Celera sequence differs from the published sequence for
this transcript
/codon_start=1
/feature="FLYBASE:FBan0002168"
/feature="FLYBASE:FBgn0017545"
/protein_id="AAF59372.1"
/feature="GI:7304341"
/translation="MAVKNKGLSKGKGGKKVVDPFSEKDWYDVKAENPMQTRQI
GKTLVNRQGORIASYLVKGRFEVSLADLQKIDIPKSFKEFLAEDVQDRNLVLCN
FHGMDLTDRKYVMVKWQTLIEAVEAKTVGYLLRVCIFGTAKDQOOSQRTCYAQ
QSOVKIRIARMTDITNEVSGADLQVLNKLALDSIAKDIKSCQRIYPLHVDYIRKV
KVLAKPRFVSKLLEHGGGKSVKVEAVSSEGAVIDRPEGEPPVOEAV"
join(30047. .30635,32975. .32989,33058. .33248,33471. .33516,
35722. .35825,43524. .43632,46598. .46685,46739. .46819,
47903. .48052,55309. .55501,55723. .55890,60156. .60226,
68154. .69555)
/feature="pan"
/feature="CT40034"
/feature="FLYBASE:FBan0017964"
/feature="FLYBASE:FBgn0019664"
/feature="FLYBASE:FBgn0019664"
/feature="pan"
/feature="CG17964"
/feature="CG17964"
/feature="102A1-102A3"
/feature="FLYBASE:FBan0017964"
/feature="FLYBASE:FBgn0019664"
join(30382. .30503,32975. .33248,33471. .33516,35722. .35825,
43524. .43632,46598. .46685,46739. .46819,47903. .48052,
55309. .55501,55723. .55890,60156. .60226,68154. .69555)
/feature="pan"
/feature="Nucleotide sequence of the Celera sequence differs
from the published sequence for this transcript."
Query Match 14.8%; Score 115.6; DB 5; Length 114334;
Best Local Similarity 51.1%; Pred. No. 1.2e-23;
Matches 271; Conservative 0; Mismatches 259; Indels 0; Gaps 0;
QY 202 actgaccagacagacagaactcattgcactggcattgctcagctgacatcacagaaat 261
DB 82504 ATTGCCCGTGAAGGCCAAGCCACTTCATGTAGCTTCGTAGTAGGAATATTAATA 82445
QY 262 gttgaattttgttgcactgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcg 321
DB 82444 ATTATGTTATTCTGCAGCATGAGCAGAGATAAATGCTCAATCTAATGACAAATATT 82385
QY 322 cctctcatattgctgcttctgctgctgctgctgctgctgctgctgctgctgctgctgctg 381
DB 82384 GCCTGTCATATTGCTGCAAAAGAACGACAGAAATATAGTCCAAAGTGTACTGGAAAT 82325
QY 382 gttgctcaagtgaatgctgcaatcaaatgctgctgctgctgctgctgctgctgctgctg 441
DB 82324 GGTGCGGAAATTAATGCTGTACCAAAAGAGGTTTACTCCACTGCATTTAGCTTCA 82265
QY 442 aaaaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 501
DB 82264 TACGGAAAGCAGAAATGCTGTTCAAAATATTACTACAAATGGAGCTTCTATTGATTT 82205
QY 502 gaccattatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 561
DB 82204 GGGAAATGATGTGACCCCATTTACACGTGGCCACGACTACAAATATCCCTCAATAG 82145
QY 562 catatctcttctactacaaagcagcagcagcagcagcagcagcagcagcagcagcagcag 621
DB 82144 GAATCTCTTTTAAATGGGTGATCCCAAAATTTATGTGCCCGAAATGCTCAATGTCT 82085
QY 622 ctactctagcctgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 681
DB 82084 ATTCACATAGCTTGCAGAAATATTTCTGAAATTCGATGCGAGCTGTGTGCAACACG 82025
QY 682 gcaagcatttcatattgagaataaagaagaagaagaagaagaagaagaagaagaagaaga 731
DB 82024 GCAGACGTTAATATTATTAGCAAGAGTGCTTCTCTCCGTTACATTGGC 81975
Search completed: August 13, 2001, 08:30:48
Job time: 2725 sec
```



```

337 GCTTCTGCTGCGCGGATGAGATTGTAAGAGCCCTTCTCGGAAAAGGTCC 388
||||: ||| :||||: ||| :||||: ||| :||||: ||| :||||: ||| :
571 AlalysTyrGlySerLeuAspValAlaLysLeuLeuGlnArgAla 572
||||: ||| :||||: ||| :||||: ||| :||||: ||| :||||: ||| :
387 TCAAGTGAATGCTGCTCAATCAAAATGGCTGCTACTCCCTTACATTTATGAG 436
||||: ||| :||||: ||| :||||: ||| :||||: ||| :||||: ||| :
597 aAlAlaAspSerAlaGlyLysAsnGlyLeuThrProLeuHisValAla 604
||||: ||| :||||: ||| :||||: ||| :||||: ||| :||||: ||| :
437 CTTTCGAAAAACAGCATGAGATCGCTGTCTGATGTTACTGGAAGCGGGGCT 486
||||: ||| :||||: ||| :||||: ||| :||||: ||| :||||: ||| :
604 laHisTyrAspAsnGlnLysValAlaLeuLeuLeuGluLysGlyAla 620
||||: ||| :||||: ||| :||||: ||| :||||: ||| :||||: ||| :
487 ANTCCAGATGCTAAGGACCATATTAGAGGTACAGCAATGCACGGGCGAC 536
||||: ||| :||||: ||| :||||: ||| :||||: ||| :||||: ||| :
621 SerProHisAlaThrAlaLysAsnGlyTyrThrProLeuHisIleAla 637
||||: ||| :||||: ||| :||||: ||| :||||: ||| :||||: ||| :
537 AGCCAAAGGGTAACTTGAAGATTCATATCTCTCTGCTACTACAAAGCAT 586
||||: ||| :||||: ||| :||||: ||| :||||: ||| :||||: ||| :
637 alYsLysAsnGlnMetGlnIleAlaSerThrLeuLeuAsnTyrGlyAla 654
||||: ||| :||||: ||| :||||: ||| :||||: ||| :||||: ||| :
587 CCACAACATCCACACACTGAGGGTAACACTCTCTCTACACTTAGCCTGT 636
||||: ||| :||||: ||| :||||: ||| :||||: ||| :||||: ||| :
654 luThrAsnIleValThrLysGlnGlyValThrProLeuHisIleAlaSer 670
||||: ||| :||||: ||| :||||: ||| :||||: ||| :||||: ||| :
637 GATGAGGAGAGGTGAAGAACAAACTGCTGTGTCCTCCCAAGGAGCAAG 686
||||: ||| :||||: ||| :||||: ||| :||||: ||| :||||: ||| :
671 GlnGluGlyHisThrAspMetValThrLeuLeuLeuAspLysGlyAla 687
||||: ||| :||||: ||| :||||: ||| :||||: ||| :||||: ||| :
687 TATTTCATTTGAGATTAAGAAGAAAAGACACCCCTGCAAGTGGCC 732
||||: ||| :||||: ||| :||||: ||| :||||: ||| :||||: ||| :
687 nIleHisMetSerThrLysSerGlyLeuThrSerLeuHisIleAla 702
||||: ||| :||||: ||| :||||: ||| :||||: ||| :||||: ||| :

seq_name: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:US-09-065-474-139

seq_documentation_block:
; Sequence 139, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 139:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

299 spMetValSerLeuLeuLeuGlyArgAsnAlaAsnValAsnLeuSerAsn 315
604 ACTGAGGGTAACACCTCTACACTTACCTGCTGATGAGGAGAGTGGA 653
316 LysSerGlyLeuThrProLeuHisLeuAlaGlnGluAspArgValAs 332
654 AGAAGCAAACTGCTGCTGCCCAAGGAGCAAGTATTTACATTGAGAATA 703
332 nValAlaGluValLeuValAsnGlnGlyAlaHisValAspAlaGlnThrL 349
704 AGAAGAAAAGACACCCCTGCAAGTG 729
349 ysMetGlyTyrThrProLeuHisVal 357

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-09-172-977-4

seq_documentation_block:
; Sequence 4, Application US/09172977
; Patent No. 5989863
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
; FILE REFERENCE: PF-0615 US
; CURRENT APPLICATION NUMBER: US/09/172,977
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO. 4
; LENGTH: 1839
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: g29491
US-09-172-977-4

alignment_scores:
Quality: 294.00 Length: 232
Ratio: 2.146 Gaps: 2
Percent Similarity: 59.052 Percent Identity: 31.897

alignment_block:
US-09-509-775-1 x US-09-172-977-4 ..
Align seg 1/1 to: US-09-172-977-4 from: 1 to: 1839

136 GCCTACAGCGGAGCTGGAAGAGTGAAGGAGAGATTCTGCGCGATAAA 185
||| :||||: ||| :||||: ||| :||||: ||| :||||: ||| :
472 AlaArgAlaGlyGlnValGluValValArg...CysLeuLeuArgAsnG1 487
186 ATCCCTGCTACTAGAACTGACGAGGACAGCAAGTCTGCTGCTGCTGCTG 235
||||: ||| :||||: ||| :||||: ||| :||||: ||| :||||: ||| :
487 yAlaLeuValAspAlaArgAlaArgGluGluGlnThrProLeuHisIleA 504
236 CATGCTCAGCTGACATACAGAATTTGTTGAATTTTGTTCGAA..... 279
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
504 laSerArgLeuGlyLysThrGluIleValGlnLeuLeuGlnHisMet 520
279 ..... 279

521 AlaHisProAspAlaAlaThrAsnGlyTyrThrProLeuHisIleSe 537
280 .....CTTGAG 286
537 rAlaArgGluGlyGlnValAspValAlaSerValLeuLeuGluAlaGlyA 554
287 TGCCAGTGAATGATAAAGACGATGACAGGTGGCTCTCTTACATTTGCG 336
||| :||||: ||| :||||: ||| :||||: ||| :||||: ||| :
554 laAlaHisSerLeuAlaThrLysLysGlyPheThrProLeuHisValAla 570

```

```
; MOLECULE TYPE: Protein
US-09-065-474-139

alignment_scores:
  Quality: 287.00      Length: 200
  Ratio: 2.142        Gaps: 2
  Percent Similarity: 67.000  Percent Identity: 36.500

alignment_block:
US-09-509-775-1 x US-09-065-474-139 ..
Align seg 1/1 to: US-09-065-474-139 from: 1 to: 352

154 CAAGAGTTGAAGGAGAGATTCTGGCC.....GA 182
   ::::: ::::: ::::: ::::: ::::: :::::
86 LysGluGlyGlnGluValAlaAlaLeuMetAspHisGlyThrAs 102
   ::::: ::::: ::::: ::::: ::::: :::::
183 TAAATCCCTGGCTACTAGAACTACAGGAGCAGGAGCACTGCATTGCAC 232
   ::::: ::::: ::::: ::::: ::::: :::::
102 pLysThrLeuLeuThrLys.....LysGlyPheThrProLeuHisL 116
   ::::: ::::: ::::: ::::: ::::: :::::
233 GGGCATGCTCAGCTGGACATACAGAAATGTTGAATTTTCTTGCACATT 282
   ::::: ::::: ::::: ::::: ::::: :::::
116 euAlaAlaLysTyrGlyAsnLeuProValAlaLysSerLeuGluArg 132
   ::::: ::::: ::::: ::::: ::::: :::::
283 GGAGTCCAGTGAATGATAAGACGATCCAGTTGGTCTCTCTTCATAT 332
   ::::: ::::: ::::: ::::: ::::: :::::
133 GlyThrProValAspLeuGluGlyLysAsnGlnValThrProLeuHisVa 149
   ::::: ::::: ::::: ::::: ::::: :::::
333 TGGCGCTTCTGCTGGCCGGGATGAGATTGTAAGCCCTTCTGGGAAAG 382
   ::::: ::::: ::::: ::::: ::::: :::::
149 lAlaAlaHisTyrAsnAsnAspLysValAlaLeuLeuLeuGluAsnG 166
   ::::: ::::: ::::: ::::: ::::: :::::
383 GTGCTCAAGTGAATGCTGCTCAATCAAAATGCTGTACTCTTACATTAT 432
   ::::: ::::: ::::: ::::: ::::: :::::
166 lYAlaSerAlaHisAlaAlaLysAsnGlyTyrThrProLeuHisIle 182
   ::::: ::::: ::::: ::::: ::::: :::::
433 GCAGCTTCGAAACAGGCATGATCGCTCATGTTTACTGGAAGCGG 482
   ::::: ::::: ::::: ::::: ::::: :::::
183 AlaAlaLysLysAsnGlnMetAspIleAlaSerThrLeuLeuHisTyrLy 199
   ::::: ::::: ::::: ::::: ::::: :::::
483 GGCTAATCCAGATGCTAAGGACCATATTAGGCTACAGCAATGCACGGG 532
   ::::: ::::: ::::: ::::: ::::: :::::
199 sAlaAsnAlaAlaGluSerLysAlaGlyPheThrProLeuHisLeuA 216
   ::::: ::::: ::::: ::::: ::::: :::::
533 CAGCAGCAAGGTAACCTTGAAGATGATTATATCTCTCTCTACTACAAA 582
   ::::: ::::: ::::: ::::: ::::: :::::
216 lAlaGlnGluGlyHisArgGluMetAlaAlaLeuLeuIleGluAsnGly 232
   ::::: ::::: ::::: ::::: ::::: :::::
583 GCATCCACAACATCCAGACACTGAGGGTAACACTCTCTACTACTAGC 632
   ::::: ::::: ::::: ::::: ::::: :::::
233 AlaLysValGlyAlaGlnAlaArgAsnGlyLeuThrProMetHisLeuCy 249
   ::::: ::::: ::::: ::::: ::::: :::::
633 CTGTGATGAGGAGAGAGTGGGAAGCAAACTGCTGGTGCCCAAGGAG 682
   ::::: ::::: ::::: ::::: ::::: :::::
249 sAlaGlnGluAspArgValSerValAlaGluLeuValLysGluAsnA 266
   ::::: ::::: ::::: ::::: ::::: :::::
683 CAAGTATTACATTGAGATAAAGCAAAAGACACCCCTTCCAAAGTGCC 732
   ::::: ::::: ::::: ::::: ::::: :::::
266 lAlaIleAspProLysThrLysAlaGlyTyrThrProLeuHisValAla 282
   ::::: ::::: ::::: ::::: ::::: :::::

seq_name: /cgn2_5/ptodata/2/laa/5B_COMB.pep.us-09-031-485-33

seq_documentation_block:
; Sequence 33, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scott
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
```

```
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,485
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429
; FILING DATE: 24-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-031-485-33

alignment_scores:
  Quality: 287.00      Length: 200
  Ratio: 2.142        Gaps: 2
  Percent Similarity: 67.000  Percent Identity: 36.500

alignment_block:
US-09-509-775-1 x US-09-031-485-33 ..
Align seg 1/1 to: US-09-031-485-33 from: 1 to: 1745

154 GAAGAGTTGAAGGAGAGATTCTGGCC.....GA 182
   ::::: ::::: ::::: ::::: ::::: :::::
521 LysGluGlyGlnGluValAlaAlaLeuMetAspHisGlyThrAs 537
   ::::: ::::: ::::: ::::: ::::: :::::
183 TAAATCCCTGGCTACTAGAACTACAGGAGCAGGAGCACTGCATTGCAC 232
   ::::: ::::: ::::: ::::: ::::: :::::
537 pLysThrLeuLeuThrLys.....LysGlyPheThrProLeuHisL 551
   ::::: ::::: ::::: ::::: ::::: :::::
233 GGGCATGCTCAGCTGGACATACAGAAATGTTGAATTTTCTTGCACATT 282
   ::::: ::::: ::::: ::::: ::::: :::::
551 euAlaAlaLysTyrGlyAsnLeuProValAlaLysSerLeuLeuGluArg 567
   ::::: ::::: ::::: ::::: ::::: :::::
283 GGAGTCCAGTGAATGATAAAGACGATGCAGTTGGTCTCTCTTTCATAT 332
   ::::: ::::: ::::: ::::: ::::: :::::
568 GlyThrProValAspLeuGluGlyLysAsnGlnValThrProLeuHisVa 584
   ::::: ::::: ::::: ::::: ::::: :::::
333 TGGCGCTTCTGCTGGCCGGGATGAGATTGTAAGCCCTTCTGGGAAAG 382
   ::::: ::::: ::::: ::::: ::::: :::::
584 lAlaAlaHisTyrAsnAsnAspLysValAlaLeuLeuLeuGluAsnG 601
   ::::: ::::: ::::: ::::: ::::: :::::
383 GTGCTCAAGTGAATGCTGCTCAATCAAAATGCTGTACTCTTACATTAT 432
   ::::: ::::: ::::: ::::: ::::: :::::
601 lYAlaSerAlaHisAlaAlaLysAsnGlyTyrThrProLeuHisIle 617
   ::::: ::::: ::::: ::::: ::::: :::::
433 GCAGCTTCGAAACAGGCATGATCGCTCATGTTTACTGGAAGCGG 482
   ::::: ::::: ::::: ::::: ::::: :::::
```

```
518 AlaAlaLysLysAsnGlnMetAspIleAlaSerThrLeuLeuHisTyrLy 634
483 GCCTAATCCATGCTAAGGACCATTTAGGCTACAGCAATGACCGGG 532
||||| :||||:||||| :||| :|||
534 sAlaAsnAlaAsnAlaGluSerLysAlaGlyPheThrProLeuHisLeuA 651
533 CAGCAGCAAGGGTAACTGAGATGATTCATATCCTCTCTACTACAAA 582
||||| :||||:||||| :||||:
651 laAlaGlnGluGlyHisArgGluMetAlaAlaLeuLeuIleGluAsnGly 667
583 GCATCCCAAAATCCCAAGACACTGAGGGTAACTCTCTACTACTTACG 632
||||| :||||:||||| :||||:
668 AlaLysValGlyAlaGlnAlaArgAsnGlyLeuThrProMetHisLeuCy 684
633 CTGTATGATGAGAGAGAGTGAAGAGCAAACTGCTGTGTGCCAAGGAG 682
684 sAlaGlnGluAspArgValSerValAlaGluGluLeuValLysGluAsnA 701
683 CAAGTATTATTCATTCAGATAAAGAGAAAGACACACCCCTGCAAGTGCC 732
||||| :||||:||||| :||||:
701 laAlaIleAspProLysThrLysAlaGlyTyrThrProLeuHisValAla 717
seq_name: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.us-08-847-429A-33

seq_documentation_block:
; Sequence 33, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429A
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-847-429A-33

alignment_scores:
Quality: 287.00 Length: 200
Ratio: 2.142 Gaps: 2
Percent Similarity: 67.000 Percent Identity: 36.500

alignment_block:
```

```
US-09-509-775-1 x US-08-847-429A-33
Align seg 1/1 to: US-08-847-429A-33 from: 1 to: 1745

154 GAAGAGTTGAAGGAGAGATATTCGGGC.....GA 182
:||||| :||||:||||| :|||
521 LysGluGlyGlnGluValAlaAlaIleLeuMetAspHisGlyThrAs 537
183 TAAATCCCTGCTACTAGAACTGACCAGGACAGCAAACTGCATTCACCT 232
||||| :||||:||||| :|||
537 PLYSThrLeuLeuThrLys.....LysGlyPheThrProLeuHisL 551
233 GGCATGCTCAGCTGAGACATACAGAAATTTGTAATTTTGTTCACACT 282
||||| :||||:||||| :||||:
551 euAlaAlaLysTyrGlyAsnLeuProValAlaLysSerLeuLeuGluArg 567
283 GGAGTCCACAGTGAATGATAAGACGATGCGAGTTGGTCTCTCTTCATAT 332
||||| :||||:||||| :||||:
568 GlyThrProValAspIleGluGlyLysAsnGlnValThrProLeuHisVa 584
333 TGGCGCTCTGCTGGCGGATGAGATTGTAAAAGCCCTTCTGGGAAAG 382
||||| :||||:||||| :||||:
584 lAlaAlaHisTyrAsnAsnAspLysValAlaLeuLeuLeuGluAsnG 601
383 GTGCTCAAGTGAATGCTCAATCAAAATGGCTGTACTCTCTTACATTAT 432
||||| :||||:||||| :||||:
601 lYAlaSerAlaHisAlaAlaLysAsnGlyTyrThrProLeuHisIle 617
433 GCAGCTTCGAAACACGCGATGAGATGCTGTCATGTTACTGGAAGCGG 482
||||| :||||:||||| :||||:
618 AlaAlaLysLysAsnGlnMetAspIleAlaSerThrLeuLeuHisTyrLy 634
483 GGCTAATCCAGATGCTAAGGACCATTTAGGCTACAGCAATGACCGGG 532
||||| :||||:||||| :||||:
634 sAlaAsnAlaAsnAlaGluSerLysAlaGlyPheThrProLeuHisLeuA 651
533 CAGCAGCAAGGGTAACTGAGATGATTCATATCCTCTCTACTACAAA 582
||||| :||||:||||| :||||:
651 laAlaGlnGluGlyHisArgGluMetAlaAlaLeuLeuIleGluAsnGly 667
583 GCATCCCAAAATCCCAAGACACTGAGGGTAACTCTCTACTACTTACG 632
||||| :||||:||||| :||||:
668 AlaLysValGlyAlaGlnAlaArgAsnGlyLeuThrProMetHisLeuCy 684
633 CTGTATGATGAGAGAGTGAAGAGCAAACTGCTGTGTGCCAAGGAG 682
||||| :||||:||||| :||||:
684 sAlaGlnGluAspArgValSerValAlaGluGluLeuValLysGluAsnA 701
683 CAAGTATTATTCATTCAGATAAAGAGAAAGACACACCCCTGCAAGTGCC 732
||||| :||||:||||| :||||:
701 laAlaIleAspProLysThrLysAlaGlyTyrThrProLeuHisValAla 717
seq_name: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.us-09-065-474-33

seq_documentation_block:
; Sequence 33, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
```

```

633 CRTGTGATGAGGAGAGACTGGAAGAAGCAAAACTGCTGGTGTGCCCAAGGAG 683
      ::::::::::::::::::::: |||||: |||||:|||||:|||||:
684 salaGlnGluAspArgValSerValAlaIaGluGluLeuValLysGluAsnA 701
      ::::::::::::::::::::: |||||: |||||:|||||:|||||:
683 CAAGTATTTCATTGAGATTAAAGCAAAAGACACACCCCTGCAACTGGCC 732
      |||:||||: :::::||||: |||||:|||||:|||||:|||||:
701 laAlaIleAspProLysThrLysAlaGlyTyrThrProLeuHisValAla 717
      ::::::::::::::::::::: |||||: |||||:|||||:|||||:

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-09-172-977-3

seq_documentation_block:
; Sequence 3, Application US/09172977
; Patent No. 5989863
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
; FILE REFERENCE: PF-0615 US
; CURRENT APPLICATION NUMBER: US/09/172,977
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 843
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE: -
; OTHER INFORMATION: g1841966
US-09-172-977-3

alignment_scores:
      Quality: 286.00      Length: 232
      Ratio: 2.103      Gaps: 2
Percent Similarity: 58.621      Percent Identity: 31.034

alignment_block:
US-09-509-775-1 x US-09-172-977-3  ..

Align seg 1/1  to: US-09-172-977-3  from: 1  to: 843

136 GCCTACAGCGGGAGCTGGAGAGATTGAAGGAGAGTATTCTGGCGCGATAA 185
      ||| ::::::::::|||: ::::: |||||: |||:
437 AlaArgAlaGlyGluValGluValArg...CysLeuLeuArgAsnG1 452
      ::::::::::::::::::::: |||||: |||||:|||||:
186 ATCCCTGGCTACTAGAACTGACGAGGACAGCAACTGCTGCACTGGG 235
      ::|||:||||: :::::||||: |||||: |||:
452 yAlaLeuValAspAlaArgAlaArgGluGluInThrProLeuHisIleA 469
      ::|||:||||: :::::||||: |||||: |||:
236 CATGCTCACTGGACATACAGAATAATTGTAATTTTGGTGCAG... 279
      || ||| |||||:|||||:|||||:|||||:
469 laSerArgLeuGlyLysThrGluIleValGlnLeuLeuGlnHisMet 485
      279 .....:|||||:|||||:|||||:|||||:
486 AlaHisProAspAlaAlaThrThrAsnGlyTyrThrProLeuHisIleSe 502
      280 .....:|||||:|||||:|||||:|||||:CTTGGAG 286
      |||:
502 rAlaArgGluGlyGlnValAspValAlaSerValLeuLeuGluAlaGlyA 519
      ::|||:
287 TGCCAGTGAATGATAAGACGATGCGAGTTGGTCTCCTCTTCATATTGG 336
      :: |||: |||||:|||||:|||||:|||||:
519 laAlaHisSerLeuAlaThrLysLysGlyPheThrProLeuHisValAla 535
      337 GCTTCTGCTGGCCGGATGAGATTGTAAGACCCCTCTGGGAAAAGGTGC 386
      |||||: |||: :::::||||: |||||: |||:
536 AlaLysTyrGlySerLeuAspValAlaIaLysLeuLeuGlnInArgArgAl 552
      387 TCAAGTGAATGTGTCAAATCAAATGGCTGTACTCCCTTCACATTATGCG 436

```

```
1 .....|
552 aAlaAspSerAlaGlyLysAsnGlyLeuThrProLeuHisValAlaA 569
437 CTTGCAAAACAGCATGAGTCCCTGCTCATCTGTTACTGGAGCGGCGCT 486
569 laHisTyrAspAsnGlnLysValAlaLeuLeuLeuGluLysGlyAla 585
487 AATCAGATGCTAAGGACCAATATGAGGCTACAGCAATGACACCGGCGC 536
586 SerProHisAlaThrAlaLysAsnGlyTyrThrProLeuHisIleAlaA 602
537 AGCCAAAGGTAACCTGAGATGATCATATCTCTGTTACTACAAAGCAT 586
602 aLysLysAsnGlnMetGlnIleAlaSerThrLeuLeuAsnTyrGlyAlaG 619
587 CCACAAACATCCAGACACTGAGGTACACTCTCTACACTTACCTGTGT 636
619 luThrAsnThrValThrLysGlnGlyValThrProLeuHisLeuAlaSer 635
637 GATGAGGAGAGTGGGAAGCAAGCAAACTGCTGTGTCCTCCAGGAGCAAG 686
636 GlnGluGlyHisThrAspMetValThrLeuValLeuGluLysGlyAlaAs 652
687 TATTTACATTCAGATAAAGAAAGAACACCCCTGCAAGTGGCC 732
652 nileHisMetSerThrLysSerGlyLeuThrSerLeuHisLeuAlaA 667

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-09-031-485-38

seq_documentation_block:
; Sequence 38, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Hesk Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,485
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429
; FILING DATE: 24-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 302 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-031-485-38
```

```
alignment_scores:
  Quality: 271.00      Length: 189
  Ratio: 2.134        Gaps: 0
  Percent Similarity: 67.196      Percent Identity: 35.450

alignment_block:
US-09-509-775-1 x US-09-031-485-38 ..
Align seg 1/1 to: US-09-031-485-38 from: 1 to: 302

166 GAGAGTATTCTGCGCCGATTAATCCCTGGCTACTAGAACTGACGACGACAG 215
||| .....|
57 GluLeuLeuLysTyrHisAlaIleGluAlaThrGluSerG1 73
216 CAGAACTGCATTGCTGCGCATGCTCAGCTGCATACAGAAATTTGTTG 265
: : : : :|
73 yLeuSerProLeuHisValAlaAlaPheMetGlyAlaIleAsnIleValI 90
266 AATTTTGTTCGAACCTGGAGTGCCTGAATGATAAAGACGATGCGAGGT 315
||| .....|
90 leYrLeuLeuGlnGlyAlaAsnAlaValAlaThrValArgGly 106
316 TGGTCTCCTTCATATGTCGGCTTCTGTCGGCGGATGAGATTGTAA 365
: : : : :|
107 GluThrProLeuHisLeuAlaArgAlaAsnGlnThrAspIleValAr 123
366 AGCCCTCTCGGAAAGGTCTCAAGTGAATGCTCAATCAAAATGGCT 415
: : : : :|
123 gValLeuValArgAsnGlyAlaGlnValAspAlaAlaArgGluLeuG 140
416 GTACTCCTTACATTTATGACCTTCGAAACACAGCATGAGTCGCTGTC 465
||| .....|
140 lnThrProLeuHisIleAlaSerArgLeuGlyAsnThrAspIleValIle 156
466 ATGTTACTGGAAGCGGGCTAATCCAGATGCTTAAGCACCATATGAGGC 515
: : : : :|
157 LeuLeuLeuGlnAlaAspAlaSerProAsnAlaAlaThrArgAspLeuTy 173
516 TACAGCAATGCACCGGCGCAGCAGCAAGGCTTAATGAAGATGATTCATA 565
||| .....|
173 rThrLeuLeuHisIleAlaAlaLysGluGlnGluValAlaAlaI 190
566 TCCTTCTGTACTACAAAGATCCACAAACATCCAAAGACACTGAGGGTAAC 615
||| .....|
190 leLeuIleAspHisGlySerAspLysThrLeuLeuThrLysLysGlyPhe 206
616 ACTCCTCTACACTTAGCTGTGTGAGGAGAGTGGGAAGACCAAACT 665
||| .....|
207 ThrProLeuHisLeuAlaAlaLysTyrGlyAsnLeuProValAlaLysLe 223
666 GCTGCTGTCCCAAGGAGCAAGTATTTACATTGAGATAAAGAGAAAAGA 715
||| .....|
223 uLeuLeuGluArgGlyThrLeuValAspIleGluGlyLysAsnGlnValT 240
716 CACCCCTGCAAGTGGCC 732
||| .....|
240 hrProLeuHisValAla 245

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-08-847-429A-38

seq_documentation_block:
; Sequence 38, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
```

```

; ADDRESS: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429A
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 302 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-847-429A-38

alignment_scores:
    Quality: 271.00      Length: 189
    Ratio: 2.134         Gaps: 0
    Percent Similarity: 67.196      Percent Identity: 35.450

alignment_block:
US-09-509-775-1 x US-08-847-429A-38 ..

Align seg 1/1 to: US-08-847-429A-38 from: 1 to: 302

166 GAGAGTATTCTGCCGCAATAATCCCTGGCTACTAGAACTGACCAGGACAG 215
||||| :||||| :||| :||| :||| :||| :||| :||| :||| :|||
57 GluLeuLeuLeuLysTyrHisAlaAlaIleGluAlaThrThrGluSerGI 73
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
216 CAGAACTGCATTGCATCGCGGCTGCTGCTGGCGGGGATGAGATTGTG 265
: :||| :||| :||| :||| :||| :||| :||| :||| :|||
73 yLeuSerProLeuHisValAlaAlaPheMetGlyAlaIleAsnIleValI 90
266 AATTTTGTGCAACTGGAGTCCAGTCCAGTGAATGATAAGACGATGAGGT 315
: :||| :||| :||| :||| :||| :||| :||| :||| :|||
90 lefyrLeuLeuGlnGlnGlyAlaAsnAlaAsnValAlaThrValArgGly 106
316 TGTCTCTCTTTCATATTGCGGCTTCTGCTGGCGGGGATGAGATTGTAA 365
: :||| :||| :||| :||| :||| :||| :||| :||| :|||
107 GluThrProLeuHisLeuAlaAlaArgAlaAsnGlnThrAspIleValar 123
366 AGCCCTCTTGGGAAAGGTGCTCAAGTGAATGCTGTCAATCAAAATGGCT 415
: :||| :||| :||| :||| :||| :||| :||| :||| :|||
123 gvalLeuValArgAsnGlyAlaGlnValAspAlaAlaAlaArgGluLeuG 140
416 GTACTCCCTTACATTATGACGCTTCGAAACACAGGATGAGATGCTGTC 465
||||| :||| :||| :||| :||| :||| :||| :||| :|||
140 lnThrProLeuHisIleAlaSerArgLeuGlyAsnThrAspIleValile 156
466 ATCTTACTGGAAGCGGGCTAATCGAGTGCCTAAGCACCATTATGAGGC 515
: :||| :||| :||| :||| :||| :||| :||| :||| :|||
157 LeuLeuLeuGlnAlaAspAlaSerProAsnAlaAlaThrArgAspLeuty 173
516 TACAGCAATGCACCGGCGAGCAGCCAGGCTAACTTGAAGATGATTCATA 565
||| :||| :||| :||| :||| :||| :||| :||| :|||
173 rThrLeuLeuHisIleAlaAlaLysGluGlyGlnGluValAlaAlaI 190
```

```

566 TCCTTCTGTACTACAAGCATCCACAACATCCAAAGACACTGAGGTTAAC 615
||||| :||| :||| :||| :||| :||| :||| :||| :|||
190 leuIleAspHisGlySerAspLysThrLeuLeuThrLysLysGlyPhe 206
616 ACTCCTCTACACTAGCTGTGATGAGGAGAGAGTGAAGAAGCAAAACT 665
||||| :||| :||| :||| :||| :||| :||| :||| :|||
207 ThrProLeuHisLeuAlaAlaLysTyrGlyAsnLeuProValAlaLysLe 223
666 GCTGCTGTCCCAAGGACAGTATTACATTGAGATTAAGAAGAAGAAGA 715
||||| :||| :||| :||| :||| :||| :||| :||| :|||
223 uLeuLeuGluArgGlyThrLeuValAspIleGluGlyLysAsnGlnVal 240
716 CACCCCTGCAAGTGCC 732
||||| :||| :||| :||| :||| :||| :||| :||| :|||
240 hrProLeuHisValAla 245

seq_name: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:US-09-065-474-38

seq_documentation_block:
; Sequence 38, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESS: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 302 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-065-474-38

alignment_scores:
    Quality: 271.00      Length: 189
    Ratio: 2.134         Gaps: 0
    Percent Similarity: 67.196      Percent Identity: 35.450

alignment_block:
US-09-509-775-1 x US-09-065-474-38 ..

Align seg 1/1 to: US-09-065-474-38 from: 1 to: 302

166 GAGAGTATTCTGCCGCAATAATCCCTGGCTACTAGAACTGACCAGGACAG 215
```

```

||| .....|||
57 GluLeuLeuLysTyrHisAlaAlaIleGluAlaThrThrGluSerGI 73
216 CAGAACTGCATTCAGCTGGGATGCTCAGCTGGACATACAGAAATTTG 265
: : : ||||| ||||| : : : |||||
73 yLeuSerProLeuHisValAlaAlaPheMetGlyAlaIleAsnIleValI 90
266 AATTTTGTTCGAACCTGGAGTGCAGTGAATGATAAGACGATGCAGGT 315
: : : ||||| ||||| : : : |||||
90 LeuTyrLeuGlnGlnGlyAlaAsnAlaAsnValAlaThrValArgGly 106
316 TGCTCTCTCTTCATATTGGCGCTTCTGCGCGGATGAGATTGTAA 365
: : : ||||| ||||| : : : |||||
107 GluThrProLeuHisLeuAlaAlaArgAlaAsnGlnThrAspIleVal 123
366 AGCCCTCTGGGAAAGGTGCTCAAGTGAATGCTCAATCAAAATGGCT 415
: : : ||||| : : : ||||| ||||| : : : |||||
123 gValLeuValArgAsnGlyAlaGlnValAspAlaAlaAlaArgGluLeu 140
416 GPACTCCCTTACATTATGACGCTTCGAAACACGACATGAGATGCTGC 465
||| ||||| ||||| ||||| : : : |||||
140 InThrProLeuHisIleAlaSerArgLeuGlyAsnThrAspIleVal 156
466 ATGTTACTGGAAGGGGCTTAATCCAGATGCTAGGACCATTTATGAGC 515
: : : ||||| : : : ||||| : : : |||||
157 LeuLeuLeuGlnAlaAspAlaSerProAsnAlaAlaThrArgAspLeu 173
516 TACAGATGCACCGGCGAGCGAGCAAGGTAACTTGAAGTATTCATA 565
||| ||||| ||||| ||||| : : : |||||
173 rThrLeuLeuHisIleAlaAlaLysGluGlnGluValAlaAlaI 190
566 TCCTCTGTACTACAAAGCATCCAAACATCCAAAGACACTGAGGTAA 615
||||| : : : ||||| : : : |||||
190 leuIleAspHisGlySerAspLysThrLeuLeuThrLysLysGlyPhe 206
616 ACTCTCTACACTTACCTGTGATGAGGAGAGAGTGGAAGAAAGTAA 665
||||| ||||| ||||| ||||| : : : |||||
207 ThrProLeuHisLeuAlaAlaLysTyrGlyAsnLeuProValAlaLys 223
566 GCTGGTGTCCCAAGGACAAGTATTACATTCAGATTAAGAAAGAAAGA 715
||| ||||| ||||| : : : |||||
223 uLeuLeuGluArgGlyThrLeuValAspIleGluGlyLysAsnGlnVal 240
716 CAGCCCTGCAAGTGCC 732
||||| ||||| |||||
240 hrProLeuHisValAla 245

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.US-09-031-485-23
seq_documentation_block:
; Sequence 23, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heskia Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Wordperfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:

```

```

APPLICATION NUMBER: US/09/031.485
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA: US/08/847.429
APPLICATION NUMBER:
FILING DATE: 24-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: HW-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 303 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-031-485-23

alignment_scores:
Quality: 262.00 Length: 218
Ratio: 1.912 Gaps: 4
Percent Similarity: 62.844 Percent Identity: 33.945

alignment_block:
US-09-509-775-1 x US-09-031-485-23
Align seg 1/1 to: US-09-031-485-23 from: 1 to: 303
136 GCCTACAGCGGAGAGCTTGAAGAGTTGAAGAGAGAGTATTCTGCCCGATAA 185
||| : : : ||| : : : |||
16 AlaHisCysGlyHisValArgValAlaLys.....LeuLeuLeuAspAr 30
186 ATCCCTGGCTACTAGAAGTACACAGGACAGCAGA.....A 220
: : : : : ||| : : : |||
30 gAsn.....AlaAspProAsnAlaArgAlaLeuAsnGlyPheT 43
221 CTGCATTGCACTGGGCGATGCTCAGCTGGACATACAGAAATTTGAATTT 270
||| ||||| ||||| : : : |||||
43 hrProLeuHisIleAlaCysLysLysAsnArgIleLysIleValGluLeu 59
271 TTGTTGCAACTTGAGTGCCAGTGAATGAATGAACAGCATGACGAGTTGGTC 320
||| : : : : : ||| : : : |||
60 LeuLeuLysTyrHisAlaAlaIleGluAlaThrThrGluSerGlyLeuSe 76
321 TCCTCTTCATATTGGGCTTCTGCTGGCGGAGTGAATTTGTAAGAGGCC 370
||| ||||| ||||| ||| : : : |||||
76 rProLeuHisValAlaAlaPheMetGlyAlaIleAsnIleValIleTyrL 93
371 TTCTGGGAAAGGTGCTCAGTGAATGCTGTCATCAATCAAAATGGCTGACT 420
||| : : : ||||| : : : |||||
93 euLeuGlnGlnGlyAlaAsnAlaAspValAlaThrValArgGlyGluThr 109
421 CCCTTACATTTATGAGCTTCCAAAACAGGATGAGATCGCTGTCATGTT 470
||| ||||| ||||| : : : |||||
110 ProLeuHisLeuAlaAlaArgAlaAsnGlnThrAspIleValArgVal 126
471 ACTGGAAGGCGGCTAATCCAGATGCTTAAGGACCATATTGAGGCTACAG 520
||| : : : ||||| : : : |||||
126 uValArgAsnGlyAlaGlnValAspAlaAlaAlaArgGluLeuGlnThrP 143
521 CAATGCACCGGCGACGCCAAGGTAACCTGAAGATGATTCATATCCTT 570
||| ||||| ||||| : : : |||||
143 roLeuHisIleAlaSerArgLeuGlyAsnThrAspIleValIleLeuLeu 159
571 CTGTACTACAAGCATCCACAACATCCAGACACTGAGGGGTAAACACTCC 620
||| : : : ||||| |||
160 LeuGlnAlaAsnAlaSerProAsnAlaAlaThrArgAspLeuTyrThrPr 176
621 TCTACACTTAGCCTGTGATGAGGAGAGTGGAGAGCAAAACACTGCTGG 670

```



```
|||||:|||||: ||| :||| ||| :|||:
176 oLeuHisIleAlaAlaLysGluGlnGluValAlaAlaIleLeu 193
671 TGTCACAGGAGCAAGTATTACATTGAGAAATAAGAAAGAACACCC 720
|||||:|||||: ||| :|||: |||: |||:
193 etAspHisGlyThrAspLysThrLeuLeuThrLysGlyPheThrPro 209
721 CTCACAGTG...GCCAAAGTGGCGCTTAAATCTCAAGAGAAATGGT 767
|||||: |||: |||: |||: |||: |||: |||: |||:
210 LeuHisLeuAlaAlaLysThrGlyAsnLeuProValAlaLysSerLeu 226
768 GGAA 771
|||
226 uGlu 227
```

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pap:US-08-847-429A-23

```
seq_documentation_block:
; Sequence 23, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heskia Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847.429A
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 303 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-847-429A-23
```

alignment_scores:
Quality: 262.00 Length: 218
Ratio: 1.912 Gaps: 4
Percent Similarity: 62.844 Percent Identity: 33.945

```
alignment_block:
US-09-509-775-1 x US-08-847-429A-23 ..
Align seg 1/1 to: US-08-847-429A-23 from: 1 to: 303
136 GCCTACAGCGGAGCTGGAAGAGTTGAAGGAGAGTATTCTGGCCGATAA 185
||||: ||| :|||: |||: |||: |||: |||: |||:
16 AlaHisCysGlyHisValArgValAlaLys.....LeuLeuLeuAsp 30
```

```
186 ATCCCTGGCTACTAGAACTGACCAGGACAGCAGA.....A 220
|||||: |||: |||: |||: |||: |||: |||: |||:
30 gAsn.....AlaAspProAsnAlaArgAlaLeuAsnGlyPheT 43
221 CTGCATTGCACTGGGCGATGCTCAGCTGGACATACAGAAATTTGTAATT 270
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
43 hrProLeuHisIleAlaCysLysLysAsnArgIleLysIleValGluLeu 59
271 TTGTTCCAACTTGGAGTGGCAGTGAATGATAAGACGATGCAGGTGGTC 320
|||||: |||: |||: |||: |||: |||: |||: |||: |||
60 LeuLeuLysThrHisAlaAlaIleGluAlaThrThrGluSerGlyLeu 76
321 TCCTCTTCATATTGCGGCTTCTGCTGGCGGGATGAGATTGTAAGACCC 370
|||||: |||: |||: |||: |||: |||: |||: |||: |||
76 rProLeuHisValAlaAlaPheMetGlyAlaIleAsnIleValIleTyrL 93
371 TTCTGGAAAGGTGCTCAAGTGAATGCTGCTCAATCAAAATGGCTGTACT 420
|||||: |||: |||: |||: |||: |||: |||: |||: |||
93 euLeuGlnGlnGlyAlaAsnAlaAspValAlaThrValArgGlyGluThr 109
421 CCCTTACATTATGCAGCTTCGAAAAACAGGCATGAGATCGCTGCTCATGT 470
|||||: |||: |||: |||: |||: |||: |||: |||: |||
110 ProLeuHisLeuAlaAlaArgAlaAsnGlnThrAspIleValArgVal 126
471 ACTGGAAGCGGGGCTAATCCAGATGCTAAGGACCATTTATGAGGCTACAG 520
|||||: |||: |||: |||: |||: |||: |||: |||: |||
126 uValArgAsnGlyAlaGlnValAspAlaAlaAlaArgGluLeuGlnThrP 143
521 CAATGCACCGGCGAGCCAGGTTAAGTGAATGATTCATATCCTT 570
|||||: |||: |||: |||: |||: |||: |||: |||: |||
143 roLeuHisIleAlaSerArgLeuGlyAsnThrAspIleValIleLeuLeu 159
571 CTGTACTACAAACCTCCACAACATCCCAAGACACTGAGGGTAACACTCC 620
||| :||| ||| ||| :||| ||| ||| :||| ||| |||
160 LeuGlnAlaAsnAlaSerProAsnAlaAlaThrArgAspLeuThrPr 176
621 TCTACACTTAGCTGTGATGAGGAGAGTGGAAAGCAAAACTGCTGG 670
|||||: |||: |||: |||: |||: |||: |||: |||: |||
176 oLeuHisIleAlaAlaLysGluGlnGluValAlaAlaIleLeu 193
671 TGTCACAGGAGCAAGTATTACATTGAGAAATAAGAAAGAACACCC 720
|||||: |||: |||: |||: |||: |||: |||: |||: |||
193 etAspHisGlyThrAspLysThrLeuLeuThrLysLysGlyPheThrPro 209
721 CTGCAAGTG...GCCAAAGTGGCGCTTAAATCTCAAGAGAAATGGT 767
|||||: |||: |||: |||: |||: |||: |||: |||: |||
210 LeuHisLeuAlaAlaLysThrGlyAsnLeuProValAlaLysSerLeu 226
768 GGAA 771
226 uGlu 227
seq_name: /cgn2_6/ptodata/2/1aa/6A_COMB.pap:US-09-065-474-23
seq_documentation_block:
; Sequence 23, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heskia Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
```



```
293 TGAATGATAAGACGATGCGAGTGGTCTCTCTCTCATATTTGGCGTCT 342
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
533 alaSnGluValAspPheGluGlyArgThrProMetHisValAlaCysGln 549
343 GCTGGCGGGATGAGATGTTAAAGCCCTCTCGGAAAGGTGCTCAAGT 392
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
550 HisGlyGlnGluAsnIleValArgThrLeuLeuArgGlyValAspVa 566
393 GAATGCTGCTCAATCAAAATGGCTACTCTTACATTTACAGCTCGA 442
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
566 lGlyLeuGlnGlyLysAspIleValArgThrLeuHisValAlaIle 583
443 AAAACAGCATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 489
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
583 lGlyHisLeuProIleValLysLeuAlaLysGlnProGlyValSer 599
490 CCAGATCTAAGGACCATTTATGAGGCTACAGCAATGCACGGGAGCAGC 539
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
600 ValAsnAlaGlnThrLeuAspGlyArgThrProLeuHisLeuAlaIle 616
540 CAAGGGTAACCTTGAAGATGATTCATATCTCTCTCTCTCTCTCTCTCT 589
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
616 naRGlyHisThrArgValAlaArgIleLeuIleAspLeuCysSerAspV 633
590 CAACATCCAGACACTGAGGTAACACTCTCTCTCTCTCTCTCTCTCTCT 639
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
633 alaSnIleCysSerLeuGlnAlaGlnThrProLeuHisValAlaIleGlu 649
640 GAGGAGAGCTGGAAGAACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 689
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
650 ThrGlyHisThrSerThrAlaArgLeuLeuHisArgGlyAlaGlyLy 666
690 TTACATTTGACAATAAGAAAGAACACCCCTGCAAGTG...GCCAAG 736
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
666 sGluAlaLeuThrSerGluGlyThrThrAlaLeuHisLeuAlaGlnA 683
737 GTGGCTGGCTTTTAATACTCAAGAGATGGTGGAA 771
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
683 snGlyHisLeuAlaThrValLysLeuLeuIleGlu 694
```

seq_name: /cgn2.6/ptodata/2/1aa/6B_COMB.pep:US-08-810-712-10

seq_documentation_block:

; Sequence 10, Application US/08810712G

; Patent No. 6160106

; GENERAL INFORMATION:

; APPLICANT: Yeda Research and Development Co. LTD

; TITLE OF INVENTION: Tumor Suppressor Genes, Proteins Encoded Thereby and

; FILE REFERENCE: Use of said Genes and Proteins

; CURRENT APPLICATION NUMBER: US/08/810,712G

; CURRENT FILING DATE: 1997-03-03

; EARLIER APPLICATION NUMBER: PCT/US94/11598

; EARLIER FILING DATE: 1994-10-12

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 10

; LENGTH: 1423

; TYPE: PRT

; ORGANISM: Homo sapiens

US-08-810-712-10

alignment_scores:

Quality: 247.00 Length: 210

Ratio: 2.111 Gaps: 1

Percent Similarity: 55.714 Percent Identity: 29.048

alignment_block:

US-09-509-775-1 x US-08-810-712-10 ..

Align seg 1/1 to: US-08-810-712-10 from: 1 to: 1423

```
205 GACCAGACAGCAGCAACTGCATTCGACTGGCATGCTCAGCTGGACATAC 254
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
401 AspLysGlyGlySerAsnAlaValTyrTrpAlaAlaArgHisGlyHisVa 417
255 AGAAATGTTGAATTTTGTGCAACTTGAGTGGCCAGTCAATGATAAAG 304
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
417 lAspThrLeuLysPheLeuSerGluAsnLysCysProLeuAspValLysA 434
305 ACGATGACAGTGTGCTCTCTCTCATATTTGCGGCTTCTGCTGGCGCGGAT 354
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
434 sPlySerGlyGluMetAlaLeuHisValAlaAlaArgTyrGlyHisAla 450
355 GAG..... 357
451 AspValAlaGlnValThrCysAlaAlaSerAlaGlnIleProIleSerAr 467
357 ..... 357
467 gThrLysGluGluThrProLeuHisCysAlaAlaTrpHisGlyTyrT 484
358 .....ATTGTAAGAGCCCTTCTGGGAAAGGTGCTCAAGTGAATGCTGTC 402
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
484 yrSerValAlaLysAlaLeuCysGluAlaGlyCysAsnValAsnIleLys 500
403 AATCAAAATGGCTGCTACTCCCTTACATTTATCCAGCTTCGAAAAACAGCCA 452
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
501 AsnArgGluGlyGluThrProLeuLeuThrAlaSerAlaArgGlyTyrHI 517
453 TGATGCTGCTGCTCATGTTACTGGAAGCGGGGCTAATCCAGATGCTTAAG 502
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
517 sAspIleValGluCysLeuAlaGluHisGlyAlaAspLeuAsnAlaCysA 534
503 ACCATTATGAGGCTACAGCAATGCACCGGCGAGCCAGCGTAACCTTG 552
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
534 sPlyAspGlyHisIleAlaLeuHisLeuAlaValArgArgCysGlnMet 550
553 AAGATGATTTCATATCCTCTCTGTACTACAAAGCATCCACAAACATCCAAGA 602
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
551 GluValIleLysThrLeuLeuSerGlnGlyCysPheValAspTyrGlnAs 567
603 CACTGAGGGTAAACACTCTCTACACTTTAGCTGTGATGAGGAGAGAGTGG 652
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
567 pArgHisGlyAsnThrProLeuHisValAlaCysLysAspGlyAsnMetP 584
653 AAGAAGCAAAACTGCTGCTGCTCCCAAGGACCAAGTATTACATTGAGAAT 702
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
584 rIleValValAlaLeuCysGluAlaAsnCysAsnLeuAspIleSerAsn 600
703 AAAGAAGAAAGACACCCCTCGCAAGTGGCC 732
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
601 LysTyrGlyArgThrProLeuHisLeuAla 610
```

THIS PAGE BLANK (USPTO)

OM of: US-09-509-775-1 to: A_Geneseq_0601.* out_format : pfs
Date: Aug 13, 2001 8:30 AM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
-MODEL=framet_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09509775/runat_13082001_074527_13026/app_query.fasta_1.848
-DB=A_Geneseq_0601 -QFMT=fastan -SUFFIX=rag -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM-ext -MINLEN=0 -MAXLEN=200000000
-USER=US09509775_@CGN1_1_51 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:

Query: US-09-509-775-1
Query length: 780
Database: A_Geneseq_0601.*
Database sequences: 412676
Database length: 60623988
Search time (sec): 74.500000

score_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
/SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT:AAW15483	+	1164.00	2268.20	226	8.3e-119
/SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT:AAW02430	+	1164.00	2268.20	226	8.3e-119
/SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT:AAW02432	+	1110.00	2161.77	231	6.9e-113
/SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT:AAW02431	+	1105.00	2151.94	231	2.4e-112
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW47022	+	298.50	547.76	1166	1.1e-23
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW66287	+	295.50	550.95	522	1.6e-23
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW66278	+	295.50	541.83	1169	2.3e-23
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW66288	+	295.50	541.83	1169	2.3e-23
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW66290	+	295.50	540.97	1169	2.4e-23
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW66294	+	295.50	539.91	1169	2.5e-23
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW66295	+	293.50	537.93	1166	3.8e-23
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW66293	+	291.50	535.83	991	5.9e-23
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW72023	+	290.50	532.03	1166	8.1e-23
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW72589	+	287.50	526.13	1166	1.7e-22
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW11616	+	287.00	538.69	352	1.1e-22
/SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT:AAW70608	+	287.00	520.59	1745	2.4e-22
/SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT:AAW76776	+	287.00	520.59	1745	2.4e-22
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW11589	+	287.00	520.59	1745	2.4e-22
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW44403	+	281.50	520.54	673	6.2e-22
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW44404	+	281.50	516.66	949	7.2e-22
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW27212	+	281.50	512.87	1327	8.4e-22
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW44402	+	281.50	512.87	1327	8.4e-22
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW66279	+	281.50	512.87	1327	8.4e-22
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW66291	+	281.00	513.20	1181	9.0e-22
/SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT:AAW70609	+	271.00	508.96	302	6.1e-21
/SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT:AAW76777	+	271.00	508.96	302	6.1e-21
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW11590	+	271.00	508.96	302	6.1e-21
/SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT:AAW3572	+	270.50	494.10	1030	1.2e-20
/SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT:AAW41378	+	270.50	494.10	1030	1.2e-20
/SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT:AAW3571	+	269.50	492.75	976	1.5e-20
/SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT:AAW41377	+	269.50	492.75	976	1.5e-20
/SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT:AAW05734	+	269.00	490.68	1074	1.8e-20
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW66286	+	265.00	486.79	756	4.2e-20
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW66285	+	265.00	486.37	784	4.2e-20
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW48574	+	265.00	484.73	907	4.5e-20
/SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT:AAW70606	+	262.00	491.22	303	5.9e-20
/SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT:AAW76774	+	262.00	491.22	303	5.9e-20
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW11587	+	262.00	491.22	303	5.9e-20
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW42288	+	260.50	475.03	978	1.5e-19
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW79154	+	259.50	475.87	763	1.7e-19

/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW69163 + 259.50 475.53 1.7e-19 786
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW76079 + 259.50 475.52 1.7e-19 787
/SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:AAW56018 + 259.50 475.52 1.7e-19 787
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW12893 + 252.50 467.92 7.8e-19 456
/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AAW27402 + 252.50 467.92 7.8e-19 456

seq_name: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT:AAW15483

seq_documentation_block:

ID AAW15483 standard; Protein; 226 AA.

XX AAW15483;

XX 17-JUN-1997 (first entry)

XX Human P28.

XX Human; proteasome; P28; diagnosis; malignant tumour.

XX Homo sapiens.

XX JP09075085-A.

XX 25-MAR-1997.

XX 13-SEP-1995; 95JP-0235052.

XX 13-SEP-1995; 95JP-0235052.

XX (SAGA) SAGAMI CHEM RES CENTRE.

XX WPI; 1997-239267/22.

XX N-PSDB; AAT66424-25.

XX Human 26S proteasome constituting component protein - useful in the diagnosis of e.g. malignant tumour

XX Claim 1; Page 6-7; 9pp; Japanese.

XX This sequence represents the human proteasome component protein P28. The protein, P28, is useful for the diagnosis and treatment of various diseases caused by proteasomes such as malignant tumour.

XX Sequence 226 AA;

alignment_scores:

Quality: 1164.00 Length: 226
Ratio: 5.150 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-509-775-1 x AAW15483

Align seg 1/1 to: AAW15483 from: 1 to: 226

97 ATGAGGGGTGTGTGTCTACCTAAATGGTGTGCAACCTGGCCTACACGG 146

||||| 1 MetGluGlyCysValSerAsnLeuMetValCysAsnLeuAlaTyrSerG1 17

147 GAGCTGGAGAGCTTCAAGCAGAGAGTATCTGGCCGCGATAAATCCCTGGCTA 196

||||| 17 yLysLeuGluGluLeuLysGluSerIleLeuAlaAspLysSerLeuAla 34

197 CTGAACTGACACAGGACAGCAGCAACCTGCACTGGGCGATGCTCAGCT 246

||||| 34 hrArgThrAspGlnAspSerArgThrAlaLeuHisTrpAlaCysSerAla 50

247 GGACATACACAGAAATGTTGAAATTTTGTGGCACTGGAGCTGCCAGTGAA 296

||||| 51 GlyHisThrGluIleValGluPheLeuLeuGluLeuGlyValProValas 67

297 TGATAAGACGATGTCAGGTGGTCTCTTCATATTGCGGCTTCTGCTG 346

```
|||||
67 nAspLysAspAspAlaGlyTpsrProLeuHisIleAlaAlaSerAlaG 84
347 GCCGGATGAGATTGTAAGCCCTTCGGGAAGAGTGCCTCAAGTGAAT 396
|||||
84 lYArgAspGluIleValLysAlaLeuLeuGlyLysGlyAlaGlnValAsn 100
397 GCTGTCAATCAAAATGGCTGTACTCCCTTACATTATGAGTTCGAAAAA 446
|||||
101 AlaValAsnGlnAsnGlyCysThrProLeuHisTyrAlaAlaSerLysAs 117
447 CAGGCATGAGATCGCTGTCATGTTACTTGAAGCGGGGCTTAATCCAGATG 496
|||||
117 nArgHisGluIleAlaValMetLeuLeuGluGlyAlaAsnProAspA 134
497 CTAAGGACCATATGAGGCTACAGCAATGCACCGGGCAGCAGCCAGGTT 546
|||||
134 lAlLysAspHisTyrGluAlaThrAlaMetHisArgAlaAlaLysGly 150
547 AACTTGAAGATCATTCATATCCTCTGCTACTACAAAGCATCCACAAACAT 596
|||||
151 AsnLeuLysMetIleHisIleLeuLeuTyrTyrLysAlaSerThrAsnI 167
597 CCAAGACACTGAGGGTAACTCCTCTACACTTAGCCTGTGATGAGGAGA 646
|||||
167 eGlnAspThrGluGlyAsnThrProLeuHisLeuAlaCysAspGluGluA 184
647 GAGTGAAGAAGCAAACTGCTGGTGTCCCAAGGAGCAAGTATTACATT 696
|||||
184 rgValGluGluAlaLysLeuLeuValSerGlnGlyAlaSerIleTyrIle 200
697 GAGATAAAGAGAAAGACACCCCTGCAAGTGGCCAAAGGTGGCTGGG 746
|||||
201 GluAsnLysGluGluLysThrProLeuGlnValAlaLysGlyIleU 217
747 TTTAATACTCAAGAGAATGGTGGAGGT 774
|||||
217 yLeuIleLeuLysArgMetValGluGly 226

seq_name: /SIDS1/gcgdata/geneseq/geneseq/AAL1999.DAT: AAY02430
seq_documentation_block:
ID AAY02430 standard; Protein; 226 AA.
AC AAY02430;
DI 14-JUL-1999 (first entry)
DE Human gankyrin protein.
KW Gankyrin; apoptosis induction; diagnosis; treatment; cancer;
KW hepatocellular carcinoma; oncogenesis mechanism.
OS Homo sapiens.
XX
XX WO9918201-A1.
XX
XX 15-APR-1999.
XX
XX 02-OCT-1998; 98WO-JP04467.
XX
XX 03-OCT-1997; 97JP-0286214.
XX
XX (FUJII) FUJITA.
XX
XX Fujita J;
XX
XX WPI; 1999-277266/23.
XX
XX N-PSDB; AAX35852.
XX
XX Gankyrin polypeptides, useful for treatment and diagnosis of
XX cancers, e.g. hepatocellular carcinoma, and study of oncogenesis
XX mechanism
```

```
XX
XX Claim 1; Page 70-71; 111pp; Japanese.
XX
CC The specification describes human, murine and rat gankyrin DNA and
CC polypeptide sequences. Gankyrin polypeptides inhibit tumorigenic
CC ability and apoptosis induction. The polypeptides and their antibodies
CC can be used in the diagnosis and treatment of cancers,
CC e.g. hepatocellular carcinoma, and study of oncogenesis mechanism.
CC The present sequence represents human gankyrin.
XX
SQ Sequence 226 AA;

alignment_scores:
Quality: 1164.00 Length: 226
Ratio: 5.150 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-509-775-1 x AAY02430 ..
Align seg 1/1 to: AAY02430 from: 1 to: 226

97 ATGAGGGGTGTGTCTTAACCTAATGTCTGCAACCTGGCTACAGCGG 146
|||||
1 MetGluGlyCysValSerAsnLeuMetValCysAsnLeuAlaTyrSerG 17

147 GAAGCTGGAAGAGTTGAGGAGAGTATTCTGGCCGATAAATCCCTGGCTA 196
|||||
17 yLysLeuGluGluLeuLysGluSerIleLeuAlaAspLysSerLeuAla 34

197 CTAGAACTGACGAGCAGCAGCACTGTCATTGCGCATGGCATGCTCAGCT 246
|||||
34 hrArgThrAspGlnAspSerArgThrAlaLeuHisTrpAlaCysSerAla 50

247 GGACATACAGAAATTTGTAATTTTGTGCACTTGGAGTGCCAGTGAA 296
|||||
51 GlyHisThrGluIleValGluPheLeuLeuGlnLeuGlyValProValAs 67

297 TGATAAAGACCATGAGGTGTGCTCCTCTTCATATTGCGGCTTCTGCTG 346
|||||
67 nAspLysAspAlaGlyTpsrProLeuHisIleAlaAlaSerAlaG 84

347 GCCGGATGAGATTGTAAGCCCTTCGGGAAGAGTGCCTCAAGTGAAT 396
|||||
84 lYArgAspGluIleValLysAlaLeuLeuGlyLysGlyAlaGlnValAsn 100

397 GCTGTCAATCAAAATGGCTGTACTCCCTTACATTATGAGTTCGAAAAA 446
|||||
101 AlaValAsnGlnAsnGlyCysThrProLeuHisTyrAlaAlaSerLysAs 117

447 CAGGCATGAGATCGCTGTCATGTTACTTGAAGCGGGGCTTAATCCAGATG 496
|||||
117 nArgHisGluIleAlaValMetLeuLeuGluGlyAlaAsnProAspA 134

497 CTAAGGACCATATGAGGCTACAGCAATGCACCGGGCAGCAGCCAGGTT 546
|||||
134 lAlLysAspHisTyrGluAlaThrAlaMetHisArgAlaAlaLysGly 150

547 AACTTGAAGATCATTCATATCCTCTGCTACTACAAAGCATCCACAAACAT 596
|||||
151 AsnLeuLysMetIleHisIleLeuLeuTyrTyrLysAlaSerThrAsnI 167

597 CCAAGACACTGAGGGTAACTCCTCTACACTTAGCCTGTGATGAGGAGA 646
|||||
167 eGlnAspThrGluGlyAsnThrProLeuHisLeuAlaCysAspGluGluA 184

647 GAGTGAAGAAGCAAACTGCTGGTGTCCCAAGGAGCAAGTATTACATT 696
|||||
184 rgValGluGluAlaLysLeuLeuValSerGlnGlyAlaSerIleTyrIle 200

697 GAGATAAAGAGAAAGACACCCCTGCAAGTGGCCAAAGGTGGCTGGG 746
|||||
```

201 GluAsnLysGluGluLysThrProLeuGlnValAlaLysGlyGlyLeuG1 217

747 TTTAATACTCAAGAGATGGTGGAGGT 774

217 yLeuIleLeuLysArgMetValGluGly 226

seq_name: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.AAY02432

seq_documentation_block:

ID AAY02432 standard; Protein; 231 AA.

XX AC AAY02432;

DT 14-JUL-1999 (first entry)

XX DE Rat gankyrin protein.

XX KW Gankyrin; apoptosis induction; diagnosis; treatment; cancer;

KW hepatocellular carcinoma; oncogenesis mechanism.

XX OS Rattus sp.

XX PN W09918201-A1.

XX PD 15-APR-1999.

XX PF 02-OCT-1998; 98WO-JP04467.

XX PR 03-OCT-1997; 97JP-0286214.

XX PA (FUJI/) FUJITA.

XX PI Fujita J;

XX DR WPI; 1999-277266/23.

XX DR N-PSDB; AAX35854.

XX KW Gankyrin polypeptides, useful for treatment and diagnosis of
PT cancers, e.g. hepatocellular carcinoma, and study of oncogenesis
PT mechanism

PS Claim 11; Page 76-78; 11pp; Japanese.

XX CC The specification describes human, murine and rat gankyrin DNA and
CC polypeptide sequences. Gankyrin polypeptides inhibit tumorigenic
CC ability and apoptosis induction. The polypeptides and their antibodies
CC can be used in the diagnosis and treatment of cancers,
CC e.g. hepatocellular carcinoma, and study of oncogenesis mechanism.
CC The present sequence represents rat gankyrin.

SQ Sequence 231 AA;

alignment_scores:

Quality: 1110.00 Length: 225

Ratio: 4.978 Gaps: 0

Percent Similarity: 99.111 Percent Identity: 95.111

alignment_block:

US-09-509-775-1 x AAY02432 ..

Align seg 1/1 to: AAY02432 from: 1 to: 231

97 ATGAGGGGTGTGTCTACCTTAATGGTGTGCAACCTGGCCTACAGCGG 146

|||||

1 MetGluGlyCysValSerAsnLeuMetValCysAsnLeuAlaTyAsnG1 17

147 GAAGCTGAAGAGTTGAAGGAGAGTATTCTGGCGGATAAATCCCTGGCTA 196

|||||

17 yLysLeuAspGluLeuLysGluSerIleLeuAlaAspLysSerLeuAla 34

197 CTAGAACTGACCGAGCAGCAGCACTGCATTGCCTGGCATGCTCAGCT 246

|||||

34 hrArgThrAspGlnAspSerArgThrAlaLeuHisTrpAlaCysSerAla 50

247 GCACATACAGAAATTTGTAATTTTGTGCAACTTGGAGTGCCAGTGAA 296

51 GlyHisThrGluIleValGluPheLeuLeuGlnLeuGlyValProValas 67

297 TGATAAAGACGATCGAGTTGGTCTCTCTTCATATTTGCGGCTTCTGCTG 346

|||||

67 nGluLysAspAlaGlyTrpSerProLeuHisIleAlaAlaSerAlaG 84

347 GCCGGGATGAGATTGTAAGCCCTTCTGGGAAAAGGTGCTCAAGTGAAT 396

|||||

84 lyArgAspGluIleValLysAlaLeuLeuIleLysGlyAlaGlnValasn 100

397 GCTGTCATCAAAATGCTGTACTCTCTTACATTATGCAGCTTCGAAAAA 446

|||||

101 AlaValasnGlnasnGlyCysThrAlaLeuHisTyrrAlaAlaSerLysas 117

447 CAGGCATGAGATCGCTCATGTTTACTGGAAAGCGGGGCTTAATCCAGATG 496

|||||

117 nArgHisGluIleAlaValMetLeuLeuGluGlyAlaAsnProaspa 134

497 CTAAGGACCATTTATGAGGCTACACCAATGCACCGGCGAGCCCAAGGT 546

|||||

134 laLysAsnHisTyrrAspAlaThrAlaMetHisArgAlaAlaAlaLysGly 150

547 AACTTGAAGATGATTCATATCTCTTCTACTACAAAGCATCCACAACAT 596

|||||

151 AsnLeuLysMetValHisIleLeuLeuPheTyrrLysAlaSerThrAsnI1 167

597 CCAAGACACTGAGGTAACACTCTCTACACTTAGCCTGTGATGAGGAGA 646

|||||

167 eGlnAspThrGluGlyAsnThrProLeuHisLeuAlaCysaspGluGluA 184

647 GAGTGGGAAGAACAAACTGCTGTGTCCTCCCAAGGAGCAAGTATTTACATT 696

|||||

184 rgValGluGluAlaLysLeuLeuValThrGlnGlyAlaSerIleTyrrIle 200

697 GAGATAAAGAAAGAAAGACACACCTCGCAAGTGGCCAAAGGTGGCTGGG 746

|||||

201 GluAsnLysGluGluLysThrProLeuGlnValAlaLysGlyGlyLeuG1 217

747 TTTAATACTCAAGAGATGGTGGAA 771

217 yLeuIleLeuLysArgIleValGlu 225

seq_name: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.AAY02431

seq_documentation_block:

ID AAY02431 standard; Protein; 231 AA.

XX AC AAY02431;

DT 14-JUL-1999 (first entry)

XX DE Mouse gankyrin protein.

XX KW Gankyrin; apoptosis induction; diagnosis; treatment; cancer;
KW hepatocellular carcinoma; oncogenesis mechanism.

XX OS Mus sp.

XX PN W09918201-A1.

XX PD 15-APR-1999.

XX PF 02-OCT-1998; 98WO-JP04467.

XX PR 03-OCT-1997; 97JP-0286214.

XX PA (FUJI/) FUJITA.

XX PI Fujita J;

XX WPI: 1999-277266/23.
DR N-PSDB; AAX35853.
XX Gankyrin polypeptides, useful for treatment and diagnosis of
PT cancers, e.g. hepatocellular carcinoma, and study of oncogenesis
PT mechanism
XX
PS Claim 6; Page 73-74; 111pp; Japanese.
XX
CC The specification describes human, murine and rat gankyrin DNA and
CC polypeptide sequences. Gankyrin polypeptides inhibit tumorigenic
CC ability and apoptosis induction. The polypeptides and their antibodies
CC can be used in the diagnosis and treatment of cancers,
CC e.g. hepatocellular carcinoma, and study of oncogenesis mechanism.
CC The present sequence represents mouse gankyrin.
XX
SQ Sequence 231 AA;

alignment_scores:
Quality: 1105.00 Length: 225
Ratio: 4.955 Gaps: 0
Percent Similarity: 99.111 Percent Identity: 93.333

alignment_block:
US-09-509-775-1 x AAY02431 ..
Align seg 1/1 to: AAY02431 from: 1 to: 231

97 ATGAGGGGTGTGCTTAACCTAATGTCACCTGGCCTACAGCGG 146
|||||
1 MetGluGlyCysValSerAsnIleMetIleCysAsnLeuAlaIleSerG1 17

147 GAGCTGGAGAGTGAAGCAGAGATCTGCGCCGATATATCCCTGGCTA 196
|||||
17 yLysLeuAspGluLeuLysGluArgIleLeuAlaAspLysSerLeuAlaI 34

197 CTAGAAGTCAACGAGCAGCAGCAATGCTGCTGCGCAGTGCAGCT 246
|||||
34 hrArgThrAspGlnAspSerArgThrAlaLeuHisTrpAlaCysSerAla 50

247 GGACATACAGAATTTGTAATTTTGTGCACTTGGAGTCCAGTGAA 296
|||||
51 GlyHisThrGluIleValGluPheLeuLeuGlnLeuGlyValProValas 67

297 TGATAAAGACGATGCGAGTGTGCTCTCTCATATTGCGGCTTCTGCTG 346
|||||
67 nAspLysAspAspAlaGlyTrpSerProLeuHisIleAlaAlaSerAlaG 84

347 GCGGGATGAGATTGTAAAGCCCTTCTGGGAAAGGTGCTCAAGTGAAT 396
|||||
84 lYArgAspGluIleValLysAlaLeuLeuValLysGlyAlaHisValasn 100

397 GGTGTCATCAATAATGCGTCTACTCCTTACATTATGCGAGTTCGAA 446
|||||
101 SerValasnGlnAsnGlyCysThrProLeuHisIleValAlaSerLysas 117

447 CAGGCATGATGCGTCTCATGTTACTGGAAGCGGGGCTAATCCAGATG 496
|||||
117 nArgHisGluIleSerValMetLeuLeuGluGlyGlyAlaasnProAspA 134

497 CTAAGGACCATATGAGGCTACAGCAATCACCGGCGCAGCAGCGGT 546
|||||
134 lAlysAspHisIleValAspAlaThrAlaMetHisArgAlaAlaLysGly 150

547 AACCTCAAGATGATTATATCTCTCTACTACAAAGCATCCACAACAT 596
|||||
151 AsnLeuLysMetValHisIleLeuLeuPheTyLysAlaSerThrAsnIle 167

597 CCAAGACATGAGGGTAACACTCCTTACACTTAGCCTGTGATGAGGAGA 646
|||||
167 eGlnAspThrGluGlyAsnThrProLeuHisLeuAlaCysAspGluGluA 184

647 GAGTGGAAAGCAAACTGCTGTGCTCCCAAGGAGCAAGTATTACATT 696
|||||
184 rgValGluGluAlaLysPheLeuValThrGlnGlyAlaSerIleTyrlle 200

697 GAGATAAAGAAAGAACACACCCCTGCAAGTGGCCAAAGTGGCCTGGG 746
|||||
201 GluAsnLysGluGluLysThrProLeuGlnValAlaLysGlyGlyLeuG1 217

747 TTTAATACTCAAGAGATGGTGGAA 771
|||||
217 yLeuIleLeuLysArgLeuAlaGlu 225

seq_name: /SIDS1/gcgdata/geneseq/AA2001.DAT:AA47022
seq_documentation_block:
ID AAB47022 standard; Protein; 1166 AA.
XX
AC AAB47022;
XX
XX 29-MAR-2001 (first entry)
DT
XX
DE Human SPANK.
XX
KW SPANK; SAM; sterile alpha motif; PAPP; insulin resistance;
KW poly adenosine diphosphate-ribose polymerase; catalytic domain;
KW ANK; ankyrin repeat; cytosol; insulin-responsive aminopeptidase;
KW IRAP; GLUT4; adipocyte; insulin signalling pathway; hyperlipidaemia;
KW glucose intolerance; atheromatous disease; atherosclerosis;
KW obesity; cardiac insufficiency; coronary insufficiency; stroke;
KW high blood pressure; non-insulin dependent diabetes; hypertension;
KW hyperuricemia; Syndrome X; muscular dystrophy; muscle atrophy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 173..209 /note= "Encoded by nucleotides 768..878"
FT FT
FT Misc-difference 327..362 /note= "Encoded by nucleotides 1230..1337"
FT FT
FT Misc-difference 375..398 /note= "Encoded by nucleotides 1374..1445"
FT FT
FT Misc-difference 482..524 /note= "Encoded by nucleotides 1695..1823"
FT FT
FT Misc-difference 641..677 /note= "Encoded by nucleotides 2172..2282"
FT FT
XX WO200077225-A1.
XX
XX 21-DEC-2000.
XX
XX 09-JUN-2000; 2000WO-US15926.
XX
XX 11-JUN-1999; 99US-0138957.
XX
XX (WHED) WHITEHEAD INST BIOMEDICAL RES.
XX (GEOH) GEN HOSPITAL CORP.
XX
XX Chi N, Lodish HF;
XX
XX WPI: 2001-091404/10.
XX N-PSDB; AAC85294.
XX
XX New insulin signalling protein SPANK, useful for reducing body mass,
XX glucose intolerance or insulin resistance and for preventing or
XX treating obesity-related and muscle-related diseases
XX
XX Claim 3; Fig 3; 65pp; English.
XX
XX This sequence represents human SPANK. The SPANK protein comprises
XX 3 domains:
XX (a) a SAM (sterile alpha motif) domain;
XX (b) a PAPP (poly adenosine diphosphate-ribose polymerase) catalytic


```

581 AAGCATCCACAACATCCAAAGACACTGAGGT..... 612
      ::::  ::  ::::  |||
182 snValAsnCysHisAlaSerAspGlyThrGlyGluTyrLysLysAspGlu 198
      ::::  ::  ::::  |||
613 .....AACACTCCTCT 623
      ::::  |||
199 LeuLeuGluSerAlaArgSerGlyAsnGluGluArgLysSerThrProLe 215
      ::::  ::  ::::  |||
624 ACACHTTAGCCTGTGATGAGGAGAGAGTGGAAGACAAACTCCTGGTGT 673
      |||||  |||  ::::  |||  ::  ::::  |||  :::
215 uHisLeuAlaLaGlyTyrAsnArgValLysIleValGlnLeuLeuLeuG 232
      |||||  |||  ::::  |||  ::  ::::  |||  :::
674 CCCAAGGAGCAGTAATTACATTGAGAAATTAAGAGAGAAAGACACCCCTG 723
      |||||  |||  ::::  |||  ::  ::::  |||  :::
232 IuHisGlyAlaAspValHisAlaLysAspLysGlyAspLeuValProLeu 248
      |||||  |||  ::::  |||  ::  ::::  |||  :::

```

```

724 CAAGTGGCC 732
      :: |||
249 HisAsnAla 251

```

seq_name: /SIDS1/gcgdata/geneseq/geneseqp/AA2001.DAT: AAB66287

```
seq_name: /SIDS1/gcgdata/geneseq/geneseq
seq_documentation_block:
ID AAB66287 standard; Protein: 522 AA.
```

ID.	AAB00287 standard;
XX	
AC	AAB66287;

AAB66287:

XX
DT
XX
XX
XX
05-APR-2001 (first entry)
Human tankyrase2 clone protein sequence SEQ ID NO: 99.
XX
DE
XX
Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder.

XX	Human tankyrase2 clone protein sequence SEQ ID NO: 99.
DE	
XX	
KW	Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW	inflammatory disorder.
XX	
OS	Homio sapiens.
XX	
PN	WO200100849-A1.
XX	
PD	04-JAN-2001.
XX	
PF	28-JUN-2000; 2000WO-US17827.
XX	
PR	29-JUN-1999; 99US-0141582.
XX	
PA	(ICOS-) ICOS CORP.
XX	
PI	Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX	
DR	WPI; 2001-102896/11.
DR	N-PSDB; AAF63925.
XX	
PT	New tankyrase2 polypeptides, useful for treating conditions mediated by
PT	poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT	inflammatory and autoimmune disorders
XX	
PS	Example 1; Page 156-157; 242pp; English.
XX	
CC	The present invention provides the protein and coding sequence for the
CC	human tankyrase2 protein. This is found in two different versions,
CC	designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
CC	polyADP-ribosylation activity and is involved in the modification of
CC	TRF1, which is a telomere-specific binding protein. The regulation of
CC	telomere length, in which TRF1 has a role, is linked to ageing and
CC	cancer. The sequences are useful in the treatment of cancers and
CC	inflammatory disorders.
XX	
SQ	Sequence 522 AA;

```

Length: 256
Gaps: 4

```

Percent Similarity: 63.281 Percent Identity: 31.641

alignment_block:

US-09-509-775-1 x AAB66287

Align seg 1/1 to: AAB66287 from: 1 to: 522

```
31 CGTAGCCGGAGCCGGCGACGTGAGCGGGCTTCTCGCGCGCAAGTAG 80
   ||| ||||| ||||| ||| |||
2 ArgileMetSerGlyArgArgCysAlaGlyGlyAlaAlaCysAlaSe 18
   ||| ||||| ||||| ||| |||
81 TTGCTGGGACAGCAAAATGAGGGGTGTCTTAACCTAATG...GTCT 127
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
18 rAlaAlaAlaGluAlaValGluProAlaAlaArgGluLeuPheGluAlaC 35
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
128 GCAACCTGGCTACAGCGGAGGTGGAAGAGTTGAAGGAGAGTATTCTG 177
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
35 ysArg.....AsnGlyAspValGluArgValLysArgLeuValThr 48
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
178 GCCGATAAATCCCTGGCTACTAGAACTGACCAGGACAGCAGAACTGCATT 227
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
49 ProGluLysValAsnSerArgAspThrAlaGlyArgLysSerThrProLe 65
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
228 GCATGGCGATGCTCAGCTGACATACAGAAATGTTGAATTTTGTTCG 277
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
65 uHisPheAlaAlaGlyPheGlyArgLysAspValValGluTyrLeuLeuG 82
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
278 RACTTGGAGTCCAGTGAATGATAAAGACGATGCAGGTTGCTCTCCTCTT 327
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
82 InAsnGlyAlaAsnValGlnAlaArgAspGlyGlyLeuIleProLeu 98
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
328 CATATGCGGGCTTCTGCGCGGATGAGATTGTAAAGCCCTTCTGGG 377
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
99 HisAsnAlaCysSerPheGlyHisAlaGluValValAsnLeuLeuLeu 115
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
378 AAAAGTGCTCAAGTGAATGTGTCAATCAAAATGGCGTGTCTCCTTAC 427
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
115 gHisGlyAlaAspProAsnAlaArgAspAsnTrpAsnTyrThrProLeu 132
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
428 ATTATGACGATCGAAAACAGGATGAGTCGCTCATGTACTGTACTGAA 477
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
132 ISgluAlaAlaIleGlyLysIleAspValCysIleValLeuLeuGln 148
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
478 GCGGGGGCTAATCCAGATGCTAAGGAC..... 504
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
149 HisGlyAlaGluProThrIleArgAsnThrAspGlyArgThrAlaLeuAs 165
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
505 .....CATTATGAGGCTA 517
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
165 pLeuAlaAspProSerAlaLysAlaValLeuThrGlyGluTyrLysLys 182
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
518 CAGCAATGACCGCGCAGCAGCAAGGTAAC...TTGAAGATGATTTCAT 564
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
182 spGluLeuLeuGluSerAlaArgSerGlyAsnGluGluLysMetMetAla 198
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
565 ATCCTCTGTACTACAAGCATCCACAACATCCAGACACTGAGGGTAA 614
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
199 LeuLeuThrProLeuAsnValAsnCysHisAlaSerAspGlyArgLysSe 215
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
615 CACTCTCTACACTTACCTGTGATCAGGAGAGAGTGAAGAGCAAAAC 664
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
215 rThrProLeuHisLeuAlaAlaGlyTyrAsnArgValLysIleValGlnL 232
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
665 TGCTGTGTGCCAAGGAGCAAGTATTTCATTGAGAATAAAGAAAG 714
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
232 euLeuLeuGlnHisGlyAlaAspValHisAlaLysAspLysGlyAspLeu 248
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
715 ACACCCCTCAAGTGCC 732
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
249 ValProLeuHisAsnAla 254
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

seq_name: /SIDS1/gcgdata/geneseq/AA2001.DAT: AAB66278

seq_documentation_block:

ID AAB66278 standard; Protein; 1169 AA.

XX AAB66278;

XX 05-APR-2001 (first entry)

XX Human tankyrase2 related protein sequence SEQ ID NO: 2.

XX Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging; inflammatory disorder.

XX Homo sapiens.

XX WO200100849-A1.

XX 04-JAN-2001.

XX 28-JUN-2000; 2000WO-US17827.

XX 29-JUN-1999; 99US-0141582.

XX (ICOS-) ICOS CORP.

XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;

XX WPI; 2001-102896/11.

XX N-PSDB; AAF63837.

XX New tankyrase2 polypeptides, useful for treating conditions mediated by poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers, inflammatory and autoimmune disorders -

XX Disclosure; Page 109-113; 242pp; English.

XX The present invention provides the protein and coding sequence for the human tankyrase2 protein. This is found in two different versions, designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-ribosylation activity and is involved in the modification of TRF1, which is a telomere-specific binding protein. The regulation of telomere length, in which TRF1 has a role, is linked to ageing and cancer. The sequences are useful in the treatment of cancers and inflammatory disorders.

XX Sequence 1169 AA;

alignment_scores:

Quality: 295.50 Length: 256

Ratio: 1.824 Gaps: 4

Percent Similarity: 63.281 Percent Identity: 31.641

alignment_block:

US-09-509-775-1 x AAB66278

Align seg 1/1 to: AAB66278 from: 1 to: 1169

31 CGTAGCCGGAGCCGGCGACGTGAGCGGGCTTCTCGCGCGCAAGTAG 80

2 ArgileMetSerGlyArgArgCysAlaGlyGlyAlaAlaCysAlaSe 18

81 TTGCTGGGACAGCAAAATGAGGGGTGTCTTAACCTAATG...GTCT 127

18 rAlaAlaAlaGluAlaValGluProAlaAlaArgGluLeuPheGluAlaC 35

128 GCAACCTGGCTACAGCGGAGGTGGAAGAGTTGAAGGAGAGTATTCTG 177

35 ysArg.....AsnGlyAspValGluArgValLysArgLeuValThr 48

178 GCCGATAAATCCCTGGCTACTAGAACTGACCAGGACAGCAAGTGCATT 227

49 ProGluLysValAsnSerArgAspThrAlaGlyArgLysSerThrProLe 65

115 aAspProAsnAlaArgAspAsnTrpAsnTyrThrProLeuHisGluAlaA 132

PR 09-JUL-1999; 99US-0350982.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2001, 07:44:23 ; Search time 32.29 Seconds
(without alignments)
424.312 Million cell updates/sec

Title: US-09-509-775-2

Perfect score: 1164

Sequence: 1 MEGCVSNLWNCNLAYSCKLE.....TPLQVAKGGLGLILKRWEG 226

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0601.*
1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1164	100.0	226	AAW15483	Human P28. Homo s
2	1164	100.0	226	AAW02430	Human gankyrin pro
3	1110	95.4	231	AAW02432	Rat gankyrin prote
4	1105	94.9	231	AAW02431	Mouse gankyrin pro
5	288.5	24.8	1166	22 AAB47022	Human SPANK. Homo
6	287	24.7	352	21 AAB11616	D. immitis ankyrin
7	287	24.7	1745	19 AAW70608	Full length ankyri
8	287	24.7	1745	19 AAW76776	D. immitis ankyrin
9	287	24.7	1745	21 AAB11589	D. immitis ankyrin
10	283.5	24.4	522	22 AAB66287	Human tankyrase2 c
11	283.5	24.4	1166	22 AAB66295	Human tankyrase2 T

12	283.5	24.4	1169	22 AAB66278	Human tankyrase2 r
13	283.5	24.4	1169	22 AAB66288	Human tankyrase2 c
14	283.5	24.4	1262	22 AAB66290	Human tankyrase2 c
15	283.5	24.4	1385	22 AAB66294	Human tankyrase2 T
16	281.5	24.2	673	21 AAY44403	Human truncated ta
17	281.5	24.2	949	21 AAY44404	Human SPANK. Mus
18	281.5	24.2	991	22 AAB47023	Mouse SPANK. Mus
19	281.5	24.2	1327	21 AAB27212	Human tankyrase I
20	281.5	24.2	1327	21 AAY44402	Human tankyrase.
21	281.5	24.2	1327	22 AAB66279	Human tankyrase2 S
22	281	24.1	1181	22 AAB66297	Drosophila tankyrase
23	280.5	24.1	1166	22 AAY72589	Human tankyrase ho
24	277.5	23.8	1166	21 AAB27211	Human tankyrase II
25	271	23.3	302	19 AAW70609	Ankyrin protein PB
26	271	23.3	302	19 AAW76777	B. malayi ankyrin
27	271	23.3	302	21 AAB11590	B. malaya ankyrin
28	270.5	23.2	1030	19 AAW53572	Human myosin light
29	270.5	23.2	1030	19 AAW41378	Human protein p164
30	269.5	23.2	976	19 AAW53571	Rat p138 protein.
31	269.5	23.2	976	19 AAW41377	Rat protein p138.
32	269	23.1	1074	20 AAY05734	Human Grb7 effecto
33	265	22.8	756	22 AAB66286	Human tankyrase2 c
34	265	22.8	784	22 AAB66285	Human tankyrase2 c
35	265	22.8	907	22 AAB48574	Human breast cance
36	262	22.5	303	19 AAW70606	Ankyrin protein fr
37	262	22.5	303	19 AAW76774	D. immitis ankyrin
38	262	22.5	303	21 AAB11587	D. immitis ankyrin
39	260.5	22.4	978	21 AAB42288	Human ORFX ORF2052
40	258.5	22.2	763	21 AAY79154	Mouse protein kina
41	258.5	22.2	786	21 AAY69163	Amino acid sequenc
42	258.5	22.2	787	21 AAY76079	Murine protein kin
43	258.5	22.2	787	22 AAB56018	Skin cell protein,
44	252.5	21.7	456	21 AAG12893	Arabidopsis thalia
45	252.5	21.7	456	21 AAG27402	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAW15483
ID AAW15483 standard; Protein; 226 AA.
XX
AC AAW15483;
XX
DT 17-JUN-1997 (first entry)
XX
DE Human P28.
XX
KW Human; proteasome; P28; diagnosis; malignant tumour.
XX
OS Homo sapiens.
XX
PN JP09075085-A.
XX
PD 25-MAR-1997.
XX
PF 13-SEP-1995; 95JP-0235052.
XX
PR 13-SEP-1995; 95JP-0235052.
XX
PA (SAGA) SAGAMI CHEM RES CENTRE.
XX
DR WPI; 1997-239267/22.
XX
N-PSDB; AAT66424-25.
XX
Human 26S proteasome constituting component protein - useful in the
diagnosis of e.g. malignant tumour
XX
Claim 1; Page 6-7; 9pp; Japanese.
XX
This sequence represents the human proteasome component protein p28.
CC The protein, p28, is useful for the diagnosis and treatment of

CC various diseases caused by proteasomes such as malignant tumour.

XX Sequence 226 AA;
 SQ
 Query Match 100.0%; Score 1164; DB 18; Length 226;
 Best Local Similarity 100.0%; Pred. No. 5e-119;
 Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MEGCVSNLMVCNLAYSGKLEELKESILADKSLATRTDQDSRTALHWACSAGHTEIVEFLL 60
 Db 1 megcvsnlmvcnlaysgkleelkesiladkslatrtdqdsrtalhwacsaghteivefll 60
 QY 61 QLGVPVNDKDDAGWSPHLHIAASAGRDEIVKALLGKGAOVNAVNGCTPLHYAASKNRHE 120
 Db 61 qlgvpvndkddagwsphlhiasagrdeivkallgkgagvnavngcctlhyaaasknrhe 120
 QY 121 IAVMLLEGANPDADKHYEATAMHRAAAGNLKMIHILLYKASTNIQDTGNTPLHLAC 180
 Db 121 iavmlleganpdadkhyeatamhraaagnlkmihiillykastniqdtgntplhlac 180
 QY 181 DEERVEEAKLLVSGASIYIENKEEKTPLQVAKGGLGLILKRMVEG 226
 Db 181 deerveeakllvsgasiyienkeektplqvakggglilkrmvveg 226
 RESULT 2
 AAY02430
 ID AAY02430 standard; Protein; 226 AA.
 XX AC AAY02430;
 XX DE Human gankyrin protein.
 XX KW Gankyrin; apoptosis induction; diagnosis; treatment; cancer;
 XX KW hepatocellular carcinoma; oncogenesis mechanism.
 XX OS Homo sapiens.
 XX PN WO9918201-A1.
 XX PD 15-APR-1999.
 XX PF 02-OCT-1998; 98WO-JP04467.
 XX PR 03-OCT-1997; 97JP-0286214.
 XX PA (FUJI/) FUJITA.
 XX PI Fujita J;
 XX DR WPI; 1999-277266/23.
 XX DR N-PSDB; AAX35852.
 XX PT Gankyrin polypeptides, useful for treatment and diagnosis of
 XX PT cancers, e.g. hepatocellular carcinoma, and study of oncogenesis
 XX PT mechanism
 XX PS Claim 1; Page 70-71; 11lpp; Japanese.
 XX CC The specification describes human, murine and rat gankyrin DNA and
 XX CC polypeptide sequences. Gankyrin polypeptides inhibit tumorigenic
 XX CC ability and apoptosis induction. The polypeptides and their antibodies
 XX CC can be used in the diagnosis and treatment of cancers,
 XX CC e.g. hepatocellular carcinoma, and study of oncogenesis mechanism.
 XX CC The present sequence represents human gankyrin.
 XX SQ Sequence 226 AA;
 Query Match 100.0%; Score 1164; DB 20; Length 226;
 Best Local Similarity 100.0%; Pred. No. 5e-119;
 Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MEGCVSNLMVCNLAYSGKLEELKESILADKSLATRTDQDSRTALHWACSAGHTEIVEFLL 60
 Db 1 megcvsnlmvcnlaysgkleelkesiladkslatrtdqdsrtalhwacsaghteivefll 60
 QY 61 QLGVPVNDKDDAGWSPHLHIAASAGRDEIVKALLGKGAOVNAVNGCTPLHYAASKNRHE 120
 Db 61 qlgvpvndkddagwsphlhiasagrdeivkallgkgagvnavngcctlhyaaasknrhe 120
 QY 121 IAVMLLEGANPDADKHYEATAMHRAAAGNLKMIHILLYKASTNIQDTGNTPLHLAC 180
 Db 121 iavmlleganpdadkhyeatamhraaagnlkmihiillykastniqdtgntplhlac 180
 QY 181 DEERVEEAKLLVSGASIYIENKEEKTPLQVAKGGLGLILKRMVEG 226
 Db 181 deerveeakllvsgasiyienkeektplqvakggglilkrmvveg 226
 RESULT 3
 AAY02432
 ID AAY02432 standard; Protein; 231 AA.
 XX AC AAY02432;
 XX DT 14-JUL-1999 (first entry)
 XX DE Rat gankyrin protein.
 XX KW Gankyrin; apoptosis induction; diagnosis; treatment; cancer;
 XX KW hepatocellular carcinoma; oncogenesis mechanism.
 XX OS Rattus sp.
 XX PN WO9918201-A1.
 XX PD 15-APR-1999.
 XX PF 02-OCT-1998; 98WO-JP04467.
 XX PR 03-OCT-1997; 97JP-0286214.
 XX PA (FUJI/) FUJITA.
 XX PI Fujita J;
 XX DR WPI; 1999-277266/23.
 XX DR N-PSDB; AAX35854.
 XX PT Gankyrin polypeptides, useful for treatment and diagnosis of
 XX PT cancers, e.g. hepatocellular carcinoma, and study of oncogenesis
 XX PT mechanism
 XX PS Claim 11; Page 76-78; 11lpp; Japanese.
 XX CC The specification describes human, murine and rat gankyrin DNA and
 XX CC polypeptide sequences. Gankyrin polypeptides inhibit tumorigenic
 XX CC ability and apoptosis induction. The polypeptides and their antibodies
 XX CC can be used in the diagnosis and treatment of cancers,
 XX CC e.g. hepatocellular carcinoma, and study of oncogenesis mechanism.
 XX CC The present sequence represents rat gankyrin.
 XX SQ Sequence 231 AA;
 Query Match 95.4%; Score 1110; DB 20; Length 231;
 Best Local Similarity 95.1%; Pred. No. 4.1e-113;
 Matches 214; Conservative 9; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MEGCVSNLMVCNLAYSGKLEELKESILADKSLATRTDQDSRTALHWACSAGHTEIVEFLL 60
 Db 1 megcvsnlmvcnlaysgkleelkesiladkslatrtdqdsrtalhwacsaghteivefll 60

Best Local Similarity 100.0%; Pred. No. 5e-119;
 Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MEGCVSNLMVCNLAYSGKLEELKESILADKSLATRTDQDSRTALHWACSAGHTEIVEFLL 60
 Db 1 megcvsnlmvcnlaysgkleelkesiladkslatrtdqdsrtalhwacsaghteivefll 60
 QY 61 QLGVPVNDKDDAGWSPHLHIAASAGRDEIVKALLGKGAOVNAVNGCTPLHYAASKNRHE 120
 Db 61 qlgvpvndkddagwsphlhiasagrdeivkallgkgagvnavngcctlhyaaasknrhe 120
 QY 121 IAVMLLEGANPDADKHYEATAMHRAAAGNLKMIHILLYKASTNIQDTGNTPLHLAC 180
 Db 121 iavmlleganpdadkhyeatamhraaagnlkmihiillykastniqdtgntplhlac 180
 QY 181 DEERVEEAKLLVSGASIYIENKEEKTPLQVAKGGLGLILKRMVEG 226
 Db 181 deerveeakllvsgasiyienkeektplqvakggglilkrmvveg 226
 RESULT 3
 AAY02432
 ID AAY02432 standard; Protein; 231 AA.
 XX AC AAY02432;
 XX DT 14-JUL-1999 (first entry)
 XX DE Rat gankyrin protein.
 XX KW Gankyrin; apoptosis induction; diagnosis; treatment; cancer;
 XX KW hepatocellular carcinoma; oncogenesis mechanism.
 XX OS Rattus sp.
 XX PN WO9918201-A1.
 XX PD 15-APR-1999.
 XX PF 02-OCT-1998; 98WO-JP04467.
 XX PR 03-OCT-1997; 97JP-0286214.
 XX PA (FUJI/) FUJITA.
 XX PI Fujita J;
 XX DR WPI; 1999-277266/23.
 XX DR N-PSDB; AAX35854.
 XX PT Gankyrin polypeptides, useful for treatment and diagnosis of
 XX PT cancers, e.g. hepatocellular carcinoma, and study of oncogenesis
 XX PT mechanism
 XX PS Claim 11; Page 76-78; 11lpp; Japanese.
 XX CC The specification describes human, murine and rat gankyrin DNA and
 XX CC polypeptide sequences. Gankyrin polypeptides inhibit tumorigenic
 XX CC ability and apoptosis induction. The polypeptides and their antibodies
 XX CC can be used in the diagnosis and treatment of cancers,
 XX CC e.g. hepatocellular carcinoma, and study of oncogenesis mechanism.
 XX CC The present sequence represents rat gankyrin.
 XX SQ Sequence 231 AA;
 Query Match 95.4%; Score 1110; DB 20; Length 231;
 Best Local Similarity 95.1%; Pred. No. 4.1e-113;
 Matches 214; Conservative 9; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MEGCVSNLMVCNLAYSGKLEELKESILADKSLATRTDQDSRTALHWACSAGHTEIVEFLL 60
 Db 1 megcvsnlmvcnlaysgkleelkesiladkslatrtdqdsrtalhwacsaghteivefll 60

[illegible]

PR	24-APR-1997;	97US-0847429.	
XX	(HESK-) HESKA CORP.		
PA			
XX	Blehm ES, Tang L;		
PI			
XX	WPI; 1998-593992/50.		
DR	N-PSDB; AAV63020.		
XX			
PT	Nucleic acids encoding ankryrins from helminth parasites - useful for		
PT	recombinant production of the proteins for use as vaccines and		
PT	treatments against helminth infection		
XX			
XX	Claim 8; Column 109-118; 84pp; English.		
XX			
CC	AAW6769-W76777 represent ankryrin proteins isolated from the helminth		
CC	parasites <i>Dirofilaria immitis</i> and <i>Brugia malayi</i> . The nucleic acids and		
CC	recombinant products are useful for the recombinant production of the		
CC	ankryrin polypeptides. These proteins can then be used as vaccines		
CC	against parasitic helminth, e.g. <i>D. immitis</i> or <i>B. malayi</i> . They can also		
CC	be used for therapy after infection, and to raise antibodies, also for		
CC	use in therapeutics, as passive immunogens, or as therapeutics against		
CC	helminths on conjugation to cytotoxic agents. The nucleic acids contained		
CC	in viruses, may also be used as viral vaccines, and the nucleic acids		
CC	themselves or in vectors may be used as genetic vaccines.		
XX			
XX	Sequence	1745 AA;	
XX			
XX	Query Match	24.7%; Score 287; DB 19; Length 1745;	
XX	Best Local Similarity	36.5%; Pred. NO. 7.7e-22;	
XX	Matches	73; Conservative 33; Mismatches 84; Indels 10; Gaps 2	
QY	20	EELKESIIA-----DKSLATRTDSDRTALHWACSGHTEIVEFLLQLGYPVNDKDDA 72	
Db	521	keggveeaaailmdhgtklltk---kgftplhlaakynlpvaksllergtvpvdiogn 577	
QY	73	GWSPLHTAASAGROEIVKALLGKGAQVNAVNGCTPLHYAASKNRHEIAVMLLEGANP 132	
Db	578	qvtplhvaahyondkvailllengasahaaakngytplhiaaknqmdiastillhykana 637	
QY	133	DAKHYEYATAMHRAAAGNKLMIHLLYYKASTNTQDTGNTPLHLACDEERVEEAKLLV 192	
Db	638	naeskagftplhlaaqeghreaalliengakvgagqarngltcmhlcagedrvsvaeelv 697	
QY	193	SGASIVYENKEETPLOVA 212	
Db	698	kenaaidpdktagytplhva 717	
XX			
XX	RESULT	9	
XX	AAB11589		
XX	ID	AAB11589 standard; Protein; 1745 AA.	
XX	AC	AAB11589;	
XX	DT	23-OCT-2000 (first entry)	
XX	DE	D. immitis ankryrin PD1ank1745.	
XX	XX		
KW	Ankryrin; parasitic helminth; filariid nematode; heartworm disease;		
KW	elephantiasis; hydrocele; vaccine; antibody; antihelminthic.		
XX	<i>Dirofilaria immitis</i> .		
OS			
XX	US06063599-A.		
PN			
XX	16-MAY-2000.		
PD			
XX			
XX	24-APR-1998;	98US-0065474.	
PF			
XX	24-APR-1997;	97US-0847429.	
XX			
XX			

PA (HESK-) HESKA CORP.

XX Blehm ES, Tang L;

XX WPI: 2000-375493/32.

DR N-PSDB; AAA58193, AAA58194, AAA58195, AAA58196.

XX New Dirofilaria and Brugia ankyrin proteins and nucleic acid encoding
PT them, useful for treating and protecting animals from diseases caused
PT by parasitic helminths, e.g. heartworm disease, elephantiasis or
PT hydrocele

XX Example 1; Column 117-128; 120pp; English.

XX The invention relates to ankyrin proteins and nucleic acids from the
CC parasitic helminths Dirofilaria immitis and Brugia malayi. It also
CC relates to antibodies raised against such ankyrin proteins and to
CC compounds that inhibit Dirofilaria or Brugia ankyrin function.
CC Dirofilaria ankyrin cDNAs were isolated from a D. immitis 48 hour
CC L3 cDNA library using PCR primers based on the sequence of the E1
CC ankyrin from Onchocerca volvulus and the Caenorhabditis elegans ankyrin
CC UNC-44 genes. Brugia ankyrin cDNAs were isolated from a B. malayi adult
CC female cDNA library using D. immitis ankyrin and C. elegans UNC-44 PCR
CC primers. Dirofilaria or Brugia ankyrin proteins and nucleic acids
CC represent novel targets for anti-helminthic vaccines and drugs. Ankyrin
CC nucleic acid molecules, proteins, vaccines and compositions are useful
CC for protecting animals, particularly dogs, from diseases caused by
CC parasitic helminths (e.g., heartworm disease, elephantiasis or
CC hydrocele), as well as for treating the infection. The ankyrin nucleic
CC acid molecules, proteins, vaccines and compositions of the invention are
CC especially useful in treating and preventing infections caused by
CC nematodes (e.g., D. immitis and B. malayi), and ascarid, capillariaid,
CC strongylid, strongyloid, trichostrongyle, or trichurid nematodes and
CC are also useful against cestodes and trematodes. The therapeutic
CC compositions may be administered to mammals, including dogs, cats,
CC humans, ferrets, horses, cattle, sheep, and other pets; economic food
CC animals; or zoo animals. The ankyrin nucleic acid molecules, proteins and
CC compounds may also be used as diagnostic reagents to detect infection by
CC parasitic helminths. Prior art anti-helminthic drugs require repeated
CC administration, which often leads to the development of resistant
CC helminth strains that no longer respond to treatment. Such drugs can
CC also cause harmful side effects in the individual being treated, and a
CC number of these drugs can only treat the symptoms of a parasitic disease,
CC being unable to prevent infection by the parasitic helminth. Elucidation
CC of D. immitis and B. malayi ankyrin protein and DNA sequences facilitates
CC the development of agents which inhibit ankyrin-mediated parasite
CC developmental and migratory pathways. Sequence AAB11589 represents full-
CC length D. immitis ankyrin, and sequences AAB11582-B11588, AAB11591-B11614
CC and AAB11616-B11629 represent D. immitis ankyrin fragments.

XX Sequence 1745 AA;

Query Match 24.7%; Score 287; DB 21; Length 1745;

Best Local Similarity 36.5%; Pred. No. 7.7e-22;

Matches 73; Conservative 33; Mismatches 84; Indels 10; Gaps 2;

Qy 20 BELKESILIA-----DKSLATRDQDSRTALHWACSGHTEIVEFLQLGVNPKDDA 72

Db 521 kegqeevaalmdhgtklltk---kgftplhaaaygnlpvaksllertgtpydiegkn 577

Qy 73 GWSPLHTAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHEIAVMLEGGANP 132

Db 578 qvtplhvaahyndnkvalillengasahaakngytpplhaaakngmdastillhykana 637

Qy 133 DAKDHYEATAMHRAAKGNLKMIIHLLYYKASTNQTEGNTPLHLACDERVEAKLLV 192

Db 638 naeskagftplhlaaqeghremaailiengakvgagarnltpmhlcaqedrsvaeely 697

Qy 193 SQGASIVYENKEETPLQVA 212

Db 698 kēnaaidpktkagtptlhva 717

RESULT 10

AAB66287

ID AAB66287 standard; Protein; 522 AA.

XX AAB66287;

XX 05-APR-2001 (first entry)

DE Human tankyrase2 clone protein sequence SEQ ID NO: 99.

KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder.

XX Homo sapiens.

XX WO200100849-A1.

XX 04-JAN-2001.

XX 28-JUN-2000; 2000WO-US17827.

XX 29-JUN-1999; 99US-0141582.

XX (ICOS-) ICOS CORP.

XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;

XX WPI: 2001-102896/11.

XX N-PSDB; AAF63925.

XX New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders
XX Example 1; Page 156-157; 242pp; English.

XX The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
CC polyADP-ribosylation activity and is involved in the modification of
CC TRF1, which is a telomere-specific binding protein. The regulation of
CC telomere length, in which TRF1 has a role, is linked to ageing and
CC cancer. The sequences are useful in the treatment of cancers and
CC inflammatory disorders.

XX Sequence 522 AA;

Query Match 24.4%; Score 283.5; DB 22; Length 522;

Best Local Similarity 32.1%; Pred. No. 3e-22;

Matches 70; Conservative 37; Mismatches 90; Indels 21; Gaps 2;

Qy 16 SGKLEELKESILADKSLATRTDQDSRTALHWACSGHTEIVEFLQLGVNPKDDAGWS 75

Db 37 ngdvervkrjltpekvnsrdtagrkstplhfaagfgrkdvveyllnganvqarddggl 96

Qy 76 PLHTAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHEIAVMLEGGANPAK 135

Db 97 plhnacsfghaeavnlllrhgadpnardnwnytpplheaakgkidvcivilqhgaetir 156

Qy 136 D-----HYTEATAMHRAAKGN-LKMIHLLYYKASTNQTEGNT 174

Db 157 ntgdrtaldladsakavltgkdkdlesarsgnekmmalltpinvnchadgrkst 216

Qy 175 PLHLACDERVEAKLLVSGASIVYENKEETPLQVA 212

Db 217 plhlaagynrvkvqlllqhgdvvhakdkgdlvplhna 254

RESULT 11

AAB66295

ID AAB66295 standard; Protein; 1166 AA.

KW inflammatory disorder.

XX

OS Homo sapiens.

XX

PN WO200100849-A1.

XX

PD 04-JAN-2001.

XX

PF 28-JUN-2000; 2000WO-US17827.

XX

PR 29-JUN-1999; 99US-0141582.

XX

PA (ICOS-) ICOS CORP.

XX

PI Christenson E, Demaggio AJ, Goldman PS, McElligott DL;

XX

DR WPI; 2001-102896/11.

XX

DR N-PSDB; AAF63926.

XX

PT New tankyrase2 polypeptides, useful for treating conditions mediated by
poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
inflammatory and autoimmune disorders -

XX

PS Example 1; Page 162-1665; 242pp; English.

XX

CC The present invention provides the protein and coding sequence for the
human tankyrase2 protein. This is found in two different versions,
designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
polyADP-ribosylation activity and is involved in the modification of
TRF1, which is a telomere-specific binding protein. The regulation of
telomere length, in which TRF1 has a role, is linked to ageing and
cancer. The sequences are useful in the treatment of cancers and
inflammatory disorders.

XX

SQ Sequence 1169 AA;

Query Match 24.4%; Score 283.5; DB 22; Length 1169;

Best Local Similarity 32.1%; Pred. No. 1e-21;

Matches 70; Conservative 37; Mismatches 90; Indels 21; Gaps 2;

QY 16 SGKLEELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFLQLGVPVNDKDDAGWS 75

Db 37 ngdverkrivtpkvnrsdrtagrkstplhfaagfgrkdvveyllnganvqardggli 96

QY 76 PLHTASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHEITAVMLLEGANPDAK 135

Db 97 plhnacsfghaevvnllrhgdpnrdnwnytplhheaaikgkidvcivilhghaeptir 156

QY 136 D-----HVEATAMHRAAAKGN-LKMIHILLYKASTNIQDTGNT 174

Db 157 ntgdrtaldladsakavltgkykdellesarsgnekemalltplnvchnasdrkst 216

QY 175 PLHACDEERVEAKLLVSOGASIYIENKEETPLQVA 212

Db 217 plhlaagynrvkivqlllghgadvhakkdgdvplhna 254

RESULT 14

AAB66290

ID AAB66290 standard; Protein; 1262 AA.

XX

AC AAB66290;

XX

DT 05-APR-2001 (first entry)

XX

DE Human tankyrase2 clone consensus protein SEQ ID NO: 107.

XX

DE Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;

KW inflammatory disorder.

XX

OS Homo sapiens.

PN WO200100849-A1.

XX

PD 04-JAN-2001.

XX

PF 28-JUN-2000; 2000WO-US17827.

XX

PR 29-JUN-1999; 99US-0141582.

XX

PA (ICOS-) ICOS CORP.

XX

PI Christenson E, Demaggio AJ, Goldman PS, McElligott DL;

XX

DR WPI; 2001-102896/11.

XX

DR N-PSDB; AAF63930.

XX

PT New tankyrase2 polypeptides, useful for treating conditions mediated by
poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
inflammatory and autoimmune disorders -

XX

PS Example 2; Page 173-176; 242pp; English.

XX

CC The present invention provides the protein and coding sequence for the
human tankyrase2 protein. This is found in two different versions,
designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
polyADP-ribosylation activity and is involved in the modification of
TRF1, which is a telomere-specific binding protein. The regulation of
telomere length, in which TRF1 has a role, is linked to ageing and
cancer. The sequences are useful in the treatment of cancers and
inflammatory disorders.

XX

SQ Sequence 1262 AA;

Query Match 24.4%; Score 283.5; DB 22; Length 1262;

Best Local Similarity 32.1%; Pred. No. 1.1e-21;

Matches 70; Conservative 37; Mismatches 90; Indels 21; Gaps 2;

QY 16 SGKLEELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFLQLGVPVNDKDDAGWS 75

Db 130 ngdverkrivtpkvnrsdrtagrkstplhfaagfgrkdvveyllnganvqardggli 189

QY 76 PLHTASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHEITAVMLLEGANPDAK 135

Db 190 plhnacsfghaevvnllrhgdpnrdnwnytplhheaaikgkidvcivilhghaeptir 249

QY 136 D-----HVEATAMHRAAAKGN-LKMIHILLYKASTNIQDTGNT 174

Db 250 ntgdrtaldladsakavltgkykdellesarsgnekemalltplnvchnasdrkst 309

QY 175 PLHACDEERVEAKLLVSOGASIYIENKEETPLQVA 212

Db 310 plhlaagynrvkivqlllghgadvhakkdgdvplhna 347

RESULT 15

AAB66294

ID AAB66294 standard; Protein; 1385 AA.

XX

AC AAB66294;

XX

DT 05-APR-2001 (first entry)

XX

DE Human tankyrase2 TANK2-LONG SEQ ID NO: 133.

XX

DE Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;

KW inflammatory disorder.

XX

OS Homo sapiens.

XX

PN WO200100849-A1.

XX

PD 04-JAN-2001.

Search completed: August 13, 2001, 07:46:05
Job time: 102 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2001, 07:44:23 ; Search time 12.84 Seconds
(without alignments)
602.939 Million cell updates/sec

Title: US-09-509-775-2

Perfect score: 1164

Sequence: 1 MEGCVSNLMVCNLAYSGKLE.....TPIQVAKGGILTKRMVEG 226

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	294	25.3	3924	1 ANK2_HUMAN	Q01484 homo sapien
2	291	25.0	228	1 YG4X_YEAST	P50086 saccharomyc
3	290	24.9	1880	1 ANK1_HUMAN	P16157 homo sapien
4	287.5	24.7	1862	1 ANK1_MOUSE	Q02357 mus musculu
5	273.5	23.5	234	1 YD57_SCHPO	Q10311 schizosacch
6	252.5	21.7	752	1 PA26_MOUSE	P97819 mus musculu
7	248	21.3	832	1 ANR3_HUMAN	P57078 homo sapien
8	247.5	21.3	751	1 PA26_RAT	P97570 rattus norv
9	247.5	21.3	1401	1 LATA_LATMA	P23631 latrodectus
10	247	21.2	1431	1 DAPK_HUMAN	P53355 homo sapien
11	236.5	20.3	439	1 AKR_ARATH	Q05753 arabidopsis
12	232	19.9	806	1 PA26_HUMAN	O60733 homo sapien
13	225	19.3	323	1 ANKH_CHRV1	Q06527 chromatium
14	224.5	19.3	898	1 KBF2_HUMAN	Q00653 homo sapien
15	217.5	18.7	452	1 ILK_MOUSE	O55222 mus musculu
16	215	18.5	656	1 FEM1_CAEEL	P17221 caenorhabdi
17	215	18.5	2703	1 NOTC_DROME	P07207 drosophila
18	214	18.4	351	1 NTC3_MOUSE	Q61982 mus musculu
19	212.5	18.3	451	1 ILK_CAVPO	P57044 cavia porce
20	211.5	18.2	452	1 ILK1_HUMAN	Q13418 homo sapien
21	211.5	18.2	452	1 ILK2_HUMAN	P57043 homo sapien
22	208.5	17.9	1964	1 NTC4_MOUSE	P31695 mus musculu
23	208	17.9	2531	1 NTC1_RAT	Q07008 rattus norv
24	205.5	17.7	2444	1 NTC1_HUMAN	P46531 homo sapien
25	202	17.4	679	1 RN5A_MOUSE	Q05921 mus musculu
26	201	17.3	446	1 BCL3_HUMAN	P20749 homo sapien
27	198	17.0	984	1 KBF1_CHICK	Q04861 gallus gall
28	197.5	17.0	2437	1 NOTC_BRARE	P46530 brachydanio
29	194.5	16.7	2524	1 NOTC_XENLA	P21783 xenopus lae
30	194	16.7	741	1 RN5A_HUMAN	Q05823 homo sapien
31	194	16.7	764	1 AKR1_YEAST	P39010 saccharomyc
32	189.5	16.3	347	1 GABC_HUMAN	Q06545 homo sapien
33	189.5	16.3	383	1 GABB_HUMAN	Q06547 homo sapien

ALIGNMENTS

RESULT 1
ANK2_HUMAN

ID ANK2_HUMAN STANDARD; PRT; 3924 AA.
AC Q01484; Q01485;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ANKYRIN 2 (BRAIN ANKYRIN) (ANKYRIN B) (ANKYRIN, NONERYTHROID).
GN ANK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RP TISSUE=Brain stem;
RC MEDLINE=94075409; PubMed=8253844;
RA Chan W., Kordeli E., Bennett V.;
RT "440-kD ankyrinB: structure of the major developmentally regulated domain and selective localization in unmyelinated axons.";
RL J. Cell Biol. 123:1463-1473(1993).
[2]
RN SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RP TISSUE=Brain stem;
RC MEDLINE=91302466; PubMed=1830053;
RA Otto E., Kunimoto M., McLaughlin T., Bennett V.;
RT "Isolation and characterization of cDNAs encoding human brain ankyrins reveal a family of alternatively spliced genes.";
RL J. Cell Biol. 114:241-253(1991).
[3]
RN REVISIONS.
RP Carpenter S.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
[4]
RN SEQUENCE OF 463-495 FROM N.A.
RC MEDLINE=92009921; PubMed=1833308;
RA Tse W.T., Menninger J.C., Yang-Feng T.L., Francke U., Sahr K.E., Lux S.E., Ward D.C., Forget B.G.;
RT "Isolation and chromosomal localization of a novel nonerythroid ankyrin gene.";
RL Genomics 10:858-866(1991).
[5]
RN FUNCTION: ANKYRINS ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTO-SKELETAL ELEMENTS. THEY BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN BAND 4.2. TO NA-K ATPASE. TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85, AND TO THE CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN. ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO THE CYTOSKELETAL DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN; THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
[6]
RN ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3; ARE PRODUCED BY ALTERNATIVE SPLICING.
[7]
RN TISSUE SPECIFICITY: PLASMA MEMBRANE OF NEURONS AS WELL AS GLIAL CELLS THROUGHOUT THE BRAIN.
[8]
RN PTM: PHOSPHORYLATED AT MULTIPLE SITES BY DIFFERENT PROTEIN KINASES AND EACH PHOSPHORYLATION EVENT REGULATES THE PROTEIN'S STRUCTURE AND FUNCTION (POTENTIAL).
[9]
RN SIMILARITY: CONTAINS 23 ANK REPEATS.

34 188 16.2 969 1 KBF1_HUMAN P19838 homo sapien
35 187.5 16.1 500 1 CACT_DROME Q03017 drosophila
36 187.5 16.1 971 1 KBF1_MOUSE P25799 mus musculu
37 187 16.1 1429 1 L112_CAEEL P14585 caenorhabdi
38 186.5 16.0 347 1 GABC_MOUSE Q00421 mus musculu
39 186.5 16.0 382 1 GABB_MOUSE Q00420 mus musculu
40 186 16.0 522 1 KBF1_RAT Q63369 rattus norv
41 183 15.7 168 1 CDN6_HUMAN P42773 homo sapien
42 182.5 15.7 757 1 HT16_HYDAT P53356 hydra atten
43 181.5 15.6 642 1 YA2A_SCHPO Q09701 schizosacch
44 179.5 15.4 2531 1 NTC1_MOUSE Q01705 mus musculu
45 178.5 15.3 414 1 GABD_MOUSE P81069 mus musculu


```
FT REPEAT 139 168 ANK 5.
FT REPEAT 173 203 ANK 6.
SQ SEQUENCE 228 AA; 25616 MW; 6B30C0DD034A6289 CRC64;

Query Match 25.0%; Score 291; DB 1; Length 228;
Best Local Similarity 35.7%; Pred. No. 6.3e-19;
Matches 75; Conservative 36; Mismatches 91; Indels 8; Gaps 4;

Qy 5 VSNLMVCNLAISGKLEELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFLQLGV 64
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 1 MSNYPHQACWENEFKVOELLKSPSLLLQDQDGRIPLHWSVSFOAHETISLLSKME 60
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

Qy 65 PVN---DKDDAGWSPLHIAASAGDEIVKALLGKGAQ--VNAVWQNGCTPLHYAASKNRH 119
   || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 61 NVNLDYDDSGWTFPHIACSVGNLVKSLYDRPLKPLDKLNKTNQGVYTLHLAVGKKWF 120
   || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

Qy 120 EIAVNLLEGGANPDQAKHYEATAMHRAAAGNLMKMHILL--YKASTNIQDTEGNTPLHL 178
   || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 121 EVSQPLIENGASVRIKDFNQIPLHRAASVGLKLIELLCGLGKSAVNWQDKQGWTPLFH 180
   || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

Qy 179 ACDEERVEAKLLYSQGASIY--IENKEEK 206
   || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 181 ALAEGHDAAVLLVEKYGAEDVLVDNKGAK 210
   || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

RESULT 3
ANK1_HUMAN STANDARD; PRT; 1880 AA.
ID ANK1_HUMAN
AC P16157;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ANKYRIN 1 (ERYTHROCYTE ANKYRIN) (ANKYRIN R) (ANKYRINS 2.1 AND 2.2).
GN ANK1 OR ANK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND VARIANTS.
RP TISSUE-Hematopoietic;
RC MEDLINE=90158830; PubMed=2137557;
RA Lux S.E., John K.M., Bennett V.;
RT "Analysis of cDNA for human erythrocyte ankyrin indicates a repeated
RT structure with homology to tissue-differentiation and cell-cycle
RT control proteins."
RL Nature 344:36-42(1990).
[2]
SEQUENCE FROM N.A.
RP MEDLINE=90175370; PubMed=1689849;
RA Lambert S., Yu H., Prchal J.T., Lawler J., Ruff P., Speicher D.,
RA Cheung M.C., Kan Y.W., Palek J.;
RT "cDNA sequence for human erythrocyte ankyrin."
RL Proc. Natl. Acad. Sci. U.S.A. 87:1730-1734(1990).
[3]
VARIANT HS ILE-462.
RP MEDLINE=96225450; PubMed=8640229;
RA Eber S.W., Gonzalez J.M., Lux M.L., Scarpa A.L., Tse W.T.,
RA Dornwell M., Herbers J., Kugler W., Oezcan R., Pekrun A.,
RA Gallagher P.G., Schroeter W., Forget B.G., Lux S.E.;
RT "Ankyrin-1 mutations are a major cause of dominant and recessive
RT hereditary spherocytosis."
RL Nat. Genet. 13:214-218(1996).
CC -1- FUNCTION: ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTOSKELETAL
CC ELEMENTS; BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN BAND 4.2, TO
CC NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85, AND TO THE
CC CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN.
CC ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO THE
CC CYTOPLASMIC DOMAIN OF THE ERYTHROCYTE ANION EXCHANGE PROTEIN;
CC THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC
CC PLASMA MEMBRANE.
```

```
CC -1- ALTERNATIVE PRODUCTS: VARIOUS ISOFORMS OF ANKYRIN ARE PRODUCED BY
CC ALTERNATIVE SPLICING. THE SEQUENCE SHOWN IS THAT OF ANKYRIN
CC VARIANT 2.1.
CC -1- P-TM: REGULATED BY PHOSPHORYLATION.
CC -1- P-TM: ACYLATED BY PALMITIC ACID GROUP(S).
CC -1- DISEASE: DEFECTS IN ANK1 ARE THE CAUSE OF DOMINANT AND RECESSIVE
CC HEREDITARY SPHEROCYTOSIS (HS).
CC -1- SIMILARITY: CONTAINS 23 ANK REPEATS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X16609; CAA34610.1; -
CC EMBL; M28880; AAA51732.1; -
CC PIR; S08275; SJHUK.
CC PIR; A35049; A35049.
CC HSP; Q00420; IAWC.
CC MIM; 182900; -
CC InterPro; IPR000488; -
CC InterPro; IPR000906; -
CC InterPro; IPR002110; -
CC Pfam; PF00791; ZU5; 1.
CC Pfam; PF00023; ank; 22.
CC Pfam; PF00531; death; 1.
CC PROSITE; PS50088; ANK_REPEAT; 20.
CC PROSITE; PS50297; ANK_REPEAT_REGION; 1.
CC PROSITE; PS50017; DEATH_DOMAIN; 1.
CC Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
CC Phosphorylation; Lipoprotein; Multigene family; Disease mutation;
CC Elliptocytosis; Polymorphism.
CC INIT_MET 0
CC DOMAIN 1 826
CC DOMAIN 827 1381
CC DOMAIN 1382 1880
CC 89 KDA DOMAIN (ANION EXCHANGE PROTEIN
CC BINDING DOMAIN).
CC 62 KDA DOMAIN (SPECTRIN BINDING
CC DOMAIN).
CC 55 KDA REGULATORY DOMAIN (REGULATES
CC THE BINDING OF ANKYRIN TO SPECTRIN
CC AND THE BAND 3 PROTEIN).
CC ANK 1.
CC ANK 2.
CC ANK 3.
CC ANK 4.
CC ANK 5.
CC ANK 6.
CC ANK 7.
CC ANK 8.
CC ANK 9.
CC ANK 10.
CC ANK 11.
CC ANK 12.
CC ANK 13.
CC ANK 14.
CC ANK 15.
CC ANK 16.
CC ANK 17.
CC ANK 18.
CC ANK 19.
CC ANK 20.
CC ANK 21.
CC ANK 22.
CC ANK 23.
CC DEATH DOMAIN.
CC MISSING (IN ISOFORM 2.2).
CC H -> D (IN ISOFORM 2.2).
CC TVEGLDPESELEVDIDYFMKSKDHTSTPNP -> ELRGS
CC GLQPDLEGRKGAQIVKRAKLRGKQ (IN A THIRD
CC ISOFORM).
CC R -> T.
CC VARIANT 20 20
```



```
Db 372 HIACKNHIRVMEILLKTGASIDAVTESGLTPLHVASPMGHLPIVKNLQRGASPNVSNV 431
Qy 171 EGNTPHLACDEEREEAKLVSOQASIIYIENKEKTPLOVAK--GGGLGLILKRMVE 225
Db 432 KVETPLHMAARAGTEVAYKLLQNKAKANAKAKDQDTPHCAARIGTGMV-KLLIE 487

RESULT 5
YD57_SCHPO STANDARD; PRT; 234 AA.
AC Q10311;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HYPOTHETICAL 25.9 KDA PROTEIN C6C3.07 IN CHROMOSOME 1.
GN SPAC6C3.07.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Devlin K.; Churcher C.M.; Barrell B.G.; Rajandream M.A.; Walsh S.V.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z69731; CA93620.1; -
CC HSSP; Q13625; 1YCS.
CC InterPro; IPR002110; -
CC Pfam; PF00023; ank; 5.
CC PROSITE; PS50088; ANK_REPEAT; 3.
CC PROSITE; PS50297; ANK_REPEAT_REGION; 1.
CC Hypothetical protein; ANK repeat; Repeat.
CC FT REPEAT 36 66 ANK 1.
CC FT REPEAT 70 100 ANK 2.
CC FT REPEAT 106 135 ANK 3.
CC FT REPEAT 140 169 ANK 4.
CC FT REPEAT 173 203 ANK 5.
CC SEQUENCE 234 AA; 25930 MW; 828F6B00A0A0B0A08E CRC64;

Query Match 23.5%; Score 273.5; DB 1; Length 234;
Best Local Similarity 34.2%; Pred. No. 2.3e-17;
Matches 68; Conservative 35; Mismatches 87; Indels 9; Gaps 5;

Qy 20 BELKESILADKSLATRTDQDSRTALHWACSAGHTEIVEFL-OLGVPVNDKDDAGWSPLH 78
Db 17 EIVEQAIQNDNSLNAVDDDKRTPLHWACSVGKVTITFLKQPNKIDKDEAGWTPLM 76
Qy 79 IAAS--AGRDEIVKALLGK-AQVNAVQNGCTPLHYAASKNRHRIAYMLLEGGANPD-- 133
Db 77 ISINRSVPDNIENRSDVDPTITRGQTCLHYAAGKGRSLIVQLCDKA--PELI 134
Qy 134 -AKDHYETAMHRAAAKNLKMHIILYYKASTNIQDTEGNTPLHACDEEREEAKLV 192
Db 135 RKKDLQGTPLHRAAAGKIQVKKYLISORAPLNTSDSYGFTPLFALAEHGPVGVELV 194
Qy 193 SOGASIIYIENKEKTPLOV 211
Db 195 RAGATLTKRDSSENHTALEV 213

RESULT 6
```

```
PA26_MOUSE
ID PA26_MOUSE STANDARD; PRT; 752 AA.
AC P97819;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE 85 KDA CALCIUM-INDEPENDENT PHOSPHOLIPASE A2 (EC 3.1.1.4) (IPLA2) (CAI-PLA2).
GN PLA2G6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DBA/2;
RX MEDLINE=97236816; PubMed=9079688;
RA Balboa M.A.; Balsinde J.; Jones S.S.; Dennis E.A.;
RT "Identity between the Ca2+-independent phospholipase A2 enzymes from p38D1 macrophages and Chinese hamster ovary cells.";
RL J. Biol. Chem. 272:8576-8580(1997).
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O -> 1-ACYLGLYCERYL-PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: CONTAINS 7 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U8624; AAB48511.1; -
CC MGD; MGI:1859152; Pla2g6.
CC HSSP; P25963; INFI.
CC InterPro; IPR002110; -
CC Pfam; PF00023; ank; 6.
CC PROSITE; PS50088; ANK_REPEAT; 4.
CC PROSITE; PS50297; ANK_REPEAT_REGION; 1.
CC Hydrolase; Lipid degradation; Repeat; ANK repeat.
CC FT REPEAT 151 181 ANK 1.
CC FT REPEAT 185 215 ANK 2.
CC FT REPEAT 219 248 ANK 3.
CC FT REPEAT 251 281 ANK 4.
CC FT REPEAT 286 312 ANK 5.
CC FT REPEAT 316 345 ANK 6.
CC FT REPEAT 349 378 ANK 7.
CC ACT_SITE 465 465 POTENTIAL.
CC SEQUENCE 752 AA; 83728 MW; 7CBE9D574741478B CRC64;

Query Match 21.7%; Score 252.5; DB 1; Length 752;
Best Local Similarity 28.9%; Pred. No. 7.3e-15;
Matches 71; Conservative 40; Mismatches 94; Indels 41; Gaps 6;

Qy 2 EGCVENLWVCNLAYSGKLEELKESILADKSLATRTDQDSRTALHWACSAGHTEIVEFLQ 61
Db 151 EGCFTPLHLACRKGDSEILVELVQYCHAQMDV---TDNKGETAHYAVQDGNPOVLQGLK 207
Qy 62 LGVP-VNKKDDAGWPLHTAASAGRDEIVKALLGKAQVNAV-----NONGC 107
Db 208 NASPGLNQVNNQGLTPLHLACKMGKQEMVRVLLLCNARCINMGPGGFTHTAMKFSQKC 267
Qy 108 -----TPLHYAASKNRHRIAYMLLEGGANPDADKHYEATAMHRA 146
Db 268 AEIIMSMSNQIHSKDPYRGASPLHWA--KNA-ENARMLLKRGCDVDSTSSSGNTALHVA 324
Qy 147 AAKGNLKMHIILYYKASTNIQDTEGNTPLHACDEEREEAKLVSOQASIIYIENKEK 206
Db 325 VMNRFDCCVMVLLTYGANAGARGEHNTPLHLAMSKDNMEMVKALIVFGAEDYDTPNDFGE 384
```

OY 207 TPLOVA 212
 Db 385 TPALIA 390

Best Local Similarity 29.6%; Pred. No. 2.1e-14;
 Matches 64; Conservative 44; Mismatches 106; Indels 2; Gaps 2;

QY 12 NLAYSGKLEELKESILADKSLATRTDQDSTALHWAACSAHTEIVEFLQLQGVFVNDKDD 71
 Db 524 HMAVERVRGVVELLARKISVNAKDEDQWALHFAAQNGDESSTRLLLEKNASVNEVDF 583

QY 72 AGWSPLHIAASAGRDEIVKALGKGAQVNAVNGCTPLHYAASKNRHEIAVMML-EGGA 130
 Db 584 EGRTPMHVACQHQENIVRILLRRGVDSVLOGKDWLPLHYAAQOGLPTVKILAKQPGV 643

QY 131 NPAKDHYEATAMHRAAKGNLKMIIHLLYYKASTNIODTGNTPLHLACDEERVEEAKL 190
 Db 644 SVNAQTLDGRTPLHLAAQGRHYVARILIDLCSDVNVCSLLAQTPLHVAETGHTSTARL 703

QY 191 LVSGQASIVTENKEEKTPLQV-AKGGILGLIKRMVE 225
 Db 704 LHRGAGKEAMTSDGYTALHLAARNHGLATVTKLIVE 739

RESULT 8
 PA26 RAT STANDARD; PRT; 751 AA.
 AC P97570;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE 85 KDA CALCIUM-INDEPENDENT PHOSPHOLIPASE A2 (EC 3.1.1.4) (IPLA2) (CAI-PLA2).
 GN PLA2G6.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Pancreatic islets;
 RX MEDLINE=97269008; PubMed=9111008;
 RA MA Z., Ramanadham S., Kempe K., Chi X.S., Ladenson J., Turk J.;
 RT "Pancreatic islets express a Ca2+-independent phospholipase A2 enzyme that contains a repeated structural homologous to the integral membrane protein binding domain of ankyrin."
 RT J. Biol. Chem. 272:11118-11127(1997).
 CC -!- FUNCTION: CATALYZES THE RELEASE OF FATTY ACIDS FROM PHOSPHOLIPIDS. IT HAS BEEN IMPLICATED IN NORMAL PHOSPHOLIPID REMODELLING, NITRIC OXIDE-INDUCED OR VASOPRESSIN-INDUCED ARACHIDONIC ACID RELEASE AND IN LEUKOTRIENE AND PROSTAGLANDIN PRODUCTION. MAY PARTICIPATE IN FAS MEDIATED APOPTOSIS AND IN REGULATING TRANSMEMBRANE ION FLUX IN GLUCOSE-STIMULATED B-CELLS.
 CC -!- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERYL-PHOSPHOCHOLINE + A FATTY ACID ANION.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -!- TISSUE SPECIFICITY: FOUND IN BRAIN, LUNG, SPLEEN, KIDNEY, LIVER, HEART AND SKELETAL MUSCLE.
 CC -!- SIMILARITY: CONTAINS 7 ANK REPEATS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; U51898; AAC53136.1;
 DR HSSP; P42773; LIHB.
 DR InterPro; IPR002110; -.
 DR Pfam; PF00023; ank; 6.
 DR PROSITE; PS50088; ANK_REPEAT; 4.
 DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
 KW Hydrolase; Lipid degradation; Repeat; ANK repeat.
 FT REPEAT 150 180
 FT REPEAT 184 214 ANK 2.

QY 207 TPLOVA 212
 Db 385 TPALIA 390

RESULT 7
 ANR3_HUMAN STANDARD; PRT; 832 AA.
 AC P57078;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE SERINE/THREONINE-PROTEIN KINASE ANKRD3 (EC 2.7.1.1-) (ANKYRIN REPEAT DOMAIN PROTEIN 3).
 GN ANKRD3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20289799; PubMed=10830953;
 RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
 RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
 RA Onki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
 RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
 RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
 RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
 RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
 RA Minoshima S., Shimizu N., Nordstiek G., Hornischer K., Brandt P.,
 RA Sharif M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
 RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
 RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
 RA Lehrscher H., Reinhardt R., Vaspo M.-L.;
 RT "The DNA sequence of human chromosome 21."
 RL Nature 405:311-319(2000).
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -!- SIMILARITY: CONTAINS 10 ANK REPEATS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; AP001743; BA95526.1; -.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
 DR PROSITE; PS50088; ANK_REPEAT; 9.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat; ANK repeat.
 FT DOMAIN 22 286 PROTEIN KINASE.
 FT REPEAT 485 514 ANK 1.
 FT REPEAT 518 547 ANK 2.
 FT REPEAT 551 580 ANK 3.
 FT REPEAT 584 613 ANK 4.
 FT REPEAT 617 647 ANK 5.
 FT REPEAT 651 680 ANK 6.
 FT REPEAT 684 713 ANK 7.
 FT REPEAT 717 746 ANK 8.
 FT REPEAT 750 780 ANK 9.
 FT REPEAT 782 811 ANK 10.
 FT NP_BIND 28 36 ATP (BY SIMILARITY).
 FT BINDING 51 51 ATP (BY SIMILARITY).
 FT ACT_SITE 143 143 BY SIMILARITY.
 SQ SEQUENCE 832 AA; 91610 MW; 5D8FFD5F04F7ECB CRC64;

Query Match 21.3%; Score 248; DB 1; Length 832;

FT REPEAT 218 247 ANK 3.
FT REPEAT 250 280 ANK 4.
FT REPEAT 285 311 ANK 5.
FT REPEAT 315 344 ANK 6.
FT REPEAT 348 377 ANK 7.
FT ACT_SITE 464 464 POTENTIAL.
SQ SEQUENCE 751 AA; 83582 MW; 393BBBADA7FCC99B CRC64;

Query Match 21.3%; Score 247.5; DB 1; Length 751;
Best Local Similarity 28.5%; Pred. No. 2e-14;
Matches 70; Conservative 39; Mismatches 96; Indels 41; Gaps 6;

QY 2 EGVSNLAVNLAISGKLEELKESILADKSLATRDQDSRTALHWACSAGTHEIVEFL-L 60
DB 150 ECTPLHLACRGDSEILVELVOYCHAQMDV---TDNGETAFHYAVOGDNPVQLQLGK 206
QY 61 QLGVVNDKDDAGNSPLHIAASAGDEIVKALLGKGAQVNAV-----NQNGC 107
DB 207 NASAGLNQVNNQGLTPLHLACQMGQEMVRVLLCNARCNINGPGGFPPIHTAMKFSQKGC 266
QY 108 -----TPLHYAKSKNRHETAVMLLEGGANPDADKHYEATAMHRA 146
DB 267 ADMIIISMSNQIHSKDPRYGASPLHWA--KNA-EMARMILLKRGCDVDSTASGNTALHYA 323
QY 147 AAKGNLKMHIILLYKASTNIODTEGNTPLHLACDEERVEEAKLLVSOGASIIYIENKEEK 206
DB 324 VTRNFDVWVLLTYGANNAGARGEGHNTPLHLAMSKDNEMVKALIVFGAEDVTPNDRGE 383

QY 207 TPLQVA 212
DB 384 TPAFIA 389

RESULT 9

QY 42 TALHWACSAGTHEIVEFLQLGVVNDKDDAGNSPLHIAASAGDEIVKALLGKGAQVNA 101
DB 732 TPLHLAVIQGRKQILSLMFDIGVNTIEQKTDKEYTPLHLAAMSKYPPELIQILLDQGSNFEA 791
QY 102 VNQNGCTPLHYAASKNRHETAVMLLEGGANPDADKHYEATAMHRAAKGNLKMHIILLY 161
DB 792 KTNNGATPLHLATFRGKSOAALLNNEVYNWRDTPDENGQMPHGAAMTGLLDVQAIIIS 851
QY 162 KAS-TNIODTEGNTPLHLACDEERVEEAKLLVSOGASIIYIENKEEKTPLOV--AKGGLGL 218
DB 852 DATVVDIEDKNSDTPNLNLAQNSHDIVIKYFIDQGDADINTRNKKGLAPLLATSKGNLDM 911
QY 219 I 219
DB 912 V 912

RESULT 10
DAPK_HUMAN
ID DAPK_HUMAN STANDARD; PRT; 1431 AA.
AC P53355;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE DEATH-ASSOCIATED PROTEIN KINASE 1 (EC 2.7.1.-) (DAP KINASE 1).
GN DAPK1 OR DAPK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95129831; PubMed=7828849;
RA Delass L.P., Feinstein E., Berissi H., Cohen O., Kimchi A.;
RT "Identification of a novel serine/threonine kinase and a novel 15-kD
RT protein as potential mediators of the gamma interferon-induced cell
RT death.";
RL Genes Dev. 9:15-30(1995).
RN [2]
RP REVISIONS TO 164-171.
RA Feinstein E.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN MEDIATING INTERFERON-GAMMA-INDUCED CELL
CC DEATH.

CC -1- PTM: AUTOPHOSPHORYLATED.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -1- SIMILARITY: CONTAINS 10 ANK REPEATS.
 CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X76104; CAA53712.1; -;
 DR HSP; Q63450; 1A06.
 DR MIM; 600831; -;
 DR InterPro; IPR000488; -;
 DR InterPro; IPR000719; -;
 DR InterPro; IPR002110; -;
 DR InterPro; IPR002290; -;
 DR Pfam; PF00023; ank; 8;
 DR Pfam; PF00531; death; 1;
 DR Pfam; PF00069; pkinase; 1;
 DR PROSITE; PS00088; ANK_REPEAT; 6;
 DR PROSITE; PS0297; ANK_REPEAT; 1;
 DR PROSITE; PS00107; DEATH_DOMAIN; 1;
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1;
 DR PROSITE; PS00101; PROTEIN_KINASE_DOM; 1;
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1;
 KW Transferase; Serine/threonine-protein kinase; Calmodulin-binding;
 KW Phosphorylation; ATP-binding; Repeat; ANK repeat; Apoptosis.
 FT DOMAIN 13 266
 FT PROTEIN KINASE.
 FT DOMAIN 267 334
 FT REPEAT 378 407
 FT REPEAT 411 440
 FT REPEAT 444 473
 FT REPEAT 478 507
 FT REPEAT 511 540
 FT REPEAT 544 573
 FT REPEAT 577 606
 FT REPEAT 610 639
 FT REPEAT 876 905
 FT REPEAT 1163 1197
 FT DOMAIN 1313 1397
 FT NP_BIND 19 27
 FT BINDING 42 42
 FT ACT_SITE 139 139
 FT MUTAGEN 42 42 K->A: LOSS OF ACTIVITY.
 SQ SEQUENCE 1431 AA; 160017 MW; 9EE84811004A15B CRC64;

Query Match 21.2%; Score 247; DB 1; Length 1431;
 Best Local Similarity 29.0%; Pred. No. 4.9e-14;
 Matches 61; Conservative 33; Mismatches 82; Indels 34; Gaps 1;
 QY 37 DQDRTALHWACSGHTEIVEFLQLGVVNDKDGWSPHLHIAASGRD----- 87
 DB 409 DKGGSNAVYWAARHGVDTLKFLSENKPLDVKDKSGEMALHVAARYGHADVAQVTCAS 468
 QY 88 -----IVKALLGKGAQVNAVQNGCTPLHYAASKNRHEIA 122
 DB 469 AQIPISRTKEETPLHCAWHGYSYSAKALCEAGCNVKNKREGETPLLTASARGYHDIV 528
 QY 123 VMLEGGANPDADKHYEATAMHRAAAAGNLMKTHILLYKASTNIQDTEGNTPLHLACDE 182
 DB 529 ECLAEHGADLNACDKDGHIALHLAVRRCQMEVITKLSQSGCFVDYQDRHGNTPLHVACKD 588
 QY 183 ERVEAKLLVSGASIIYENKEKTPLOVA 212
 DB 589 GNMPIVVALCEANCLDINSRYGRTPLHLA 618
 RESULT 11

AKR_ARATH
 ID AKR_ARATH STANDARD; PRT; 439 AA.
 AC Q05753;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE ANKYRIN REPEAT PROTEIN (AKRP).
 GN AKR.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. C24; TISSUE=Leaf;
 RX MEDLINE=93104681; PubMed=1281700;
 RA Zhang H., Scheirer D.C., Fowle W.H., Goodman H.M.;
 RT "Expression of antisense or sense RNA of an ankyrin repeat-containing
 gene blocks chloroplast differentiation in Arabidopsis.";
 RL Plant Cell 4:1575-1588(1992).
 CC -1- FUNCTION: MAY PLAY A ROLE IN THE TEMPORAL AND SPATIAL REGULATION
 CC OF CHLOROPLAST DEVELOPMENT FROM PROPLASTID.
 CC -1- DEVELOPMENTAL STAGE: HIGHEST EXPRESSION OCCURS IN TWO-WEEK-OLD
 CC PLANTS AND DECLINES AS PLANTS DEVELOP FURTHER.
 CC -1- SIMILARITY: CONTAINS 4 ANK REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M82883; AAA32812.1; -;
 DR PIR; JQ1729; JQ1729.
 DR HSP; Q13625; 1YCS.
 DR InterPro; IPR002110; -;
 DR Pfam; PF00023; ank; 4;
 DR PROSITE; PS50088; ANK_REPEAT; 3;
 DR PROSITE; PS50297; ANK_REPEAT; 1;
 KW Cytoskeleton; Repeat; ANK repeat; Multigene family.
 FT REPEAT 288 317
 FT REPEAT 321 350
 FT REPEAT 354 383
 FT REPEAT 387 416
 FT SEQUENCE 439 AA; 49150 MW; C371A90028B25BF3 CRC64;
 Query Match 20.3%; Score 236.5; DB 1; Length 439;
 Best Local Similarity 37.3%; Pred. No. 1e-13;
 Matches 60; Conservative 19; Mismatches 73; Indels 9; Gaps 2;
 QY 57 EFLL-----QLGVVNDKDGWSPHLHIAASGRDEIVKALLGKGAQVNAVQNGCTPLH 111
 DB 272 KFMLNSRNPDLAVATSKK---WLPLHTLAACGEFYLDVSLKHLNDINATDVGGTLALH 327
 QY 112 YAASKNRHETAVMLLEGGANPDADKHYEATAMHRAAAAGNLMKTHILLYKASTNIQDTE 171
 DB 328 RAIIGKQATNYLLRSANFVLDDEGATLMHVAVTASAPTIKLLLYNADINAQDRD 387
 QY 172 GNTPLHLACDEERVEEAKLLVSGASIIYENKEKTPLOVA 212
 DB 388 GWTPLHVAVQARRSDIVKLLLIKADIEVKNKDGLTPLGEA 428
 RESULT 12
 ID PA26_HUMAN STANDARD; PRT; 806 AA.
 AC Q06733; Q75645; Q9Y671; Q9UIT0; Q9UG29;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)

Search completed: August 13, 2001, 07:44:50
Job time: 27 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2001, 07:44:23 ; Search time 22.54 Seconds
(without alignments)
763.773 Million cell updates/sec

Title: US-09-509-775-2
Perfect score: 1164
Sequence: 1 MEGCVSNLMVCNLAYSGKLE.....TPLQVAKGGLGLILKRWVEG 226

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	338	29.0	247	2 DB4448	probable ankyrin [
2	312.5	26.8	1765	2 T42714	ankyrin 3, splice
3	312.5	26.8	1940	2 T42715	ankyrin 3, splice
4	312.5	26.8	1943	2 T42713	ankyrin 3, splice
5	312.5	26.8	1961	2 T42716	ankyrin 3, splice
6	310	26.6	237	2 T50984	related to 26s pro
7	310	26.6	4377	2 A55575	ankyrin 3, long sp
8	305	26.2	636	2 T33631	hypothetical prote
9	294	25.3	3924	2 S37431	ankyrin 2, neuro
10	291	25.0	228	2 S57697	hypothetical prote
11	290	24.9	1856	2 B35049	ankyrin 1, erythro
12	290	24.9	1880	2 A35049	ankyrin 1, erythro
13	290	24.9	1881	1 SHUK	ankyrin 1, erythro
14	288.5	24.8	1786	2 A57282	ankyrin-related pr
15	288.5	24.8	1815	2 T15346	elegans ankyrin-re
16	288.5	24.8	1867	2 T15344	ankyrin-related un
17	288.5	24.8	2039	2 T15347	ankyrin-related un
18	287.5	24.7	1848	2 S37771	ankyrin, erythrocy
19	287.5	24.7	1862	2 T49502	ankyrin - mouse
20	276.5	23.8	1549	2 T13940	ankyrin - fruit fl
21	273.5	23.5	234	2 T39032	hypothetical ankyr
22	270.5	23.2	815	2 JG0197	myosin-light-chain
23	269.5	23.2	658	2 S68418	protein phosphatas
24	268.5	23.1	1004	2 A55142	myosin-light-chain
25	260.5	22.4	1062	2 T30255	inversin - mouse
26	248	21.3	397	2 T46445	hypothetical prote
27	248	21.3	1062	2 T14151	inv protein - mous
28	247.5	21.3	1401	2 S11527	alpha-latrotoxin p
29	247	21.2	1423	1. I37275	death-associated p

30	243.5	20.9	1411	2 S30355	alpha-latroinsecto
31	239.5	20.6	368	2 T18184	ankyrin repeat pro
32	239	20.5	319	2 A57291	cytokine inducible
33	239	20.5	557	2 T46507	hypothetical prote
34	237.5	20.4	791	2 T42691	hypothetical prote
35	236.5	20.3	439	2 JQ1729	ankyrin-repeat pro
36	233	20.0	934	1 H71274	probable ankyrin -
37	232	19.9	851	2 T12503	hypothetical prote
38	232	19.9	1398	2 T21884	hypothetical prote
39	230	19.8	1031	2 T43458	hypothetical prote
40	228	19.6	2352	2 T30201	Notch homolog prot
41	226.5	19.5	1001	2 S30385	G9a protein - huma
42	225.5	19.4	1188	2 T19552	hypothetical prote
43	225	19.3	323	2 B47169	ankyrin-like repea
44	224.5	19.3	900	2 A42024	transcription fact
45	224.5	19.3	933	2 S17233	transcription fact

ALIGNMENTS

RESULT 1

DB4448
probable ankyrin [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: DB4448
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: DB4448
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-247 <STO>
A:Cross-references: GB:AE002093; NID:94335756; PIDN:AAD17433.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g03430
A:Map position: 2

Query Match 29.0%; Score 338; DB 2; Length 247;

Best Local Similarity 41.1%; Pred. No. 6.1e-22;

Matches 90; Conservative 32; Mismatches 81; Indels 16; Gaps 6;

Qy 22 LKESILADKSLATRDQDSRTALHWACSAGHTEIVEFLL---QLGVPVNDKDDAGMSPLH 78

Db 31 LSEEQLS-KSLNFR-NEDGRSLHVAASFQSVKLLSSSDSEAKTVINSKDDGMAPLH 88

Qy 79 IAAAGRDEIVKALGKGAOVNAVNGCTPLHYAASNHRHIAVVLLEGGANPDADKHV 138

Db 89 SAAISGNAELVEVLLTRGADVNAKNNGGRTALHYAASGRLEIAQLLLTHGAKINITDKV 148

Qy 139 EATAMHRAAGNKLMIHLLYYKASTNIDQTEGNTPL---HLACDEERV-----EAKLL 191

Db 149 GCTPLHRAASGVKLEVFLEEAGAEIDATDKMGOTALMHSVICDDKQLKVSMDQVAFLL 208

Qy 192 VSQASIVIEKKEETPLQVAKGGLGLIL---KRWEG 226

Db 209 IRHGADVDEKGYTVLGRATNEFRPALIDAAKAMLEG 247

RESULT 2

T42714
ankyrin 3, splice form 2 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C:Accession: T42714
R:Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Vialamas, M.; Turtzo
J. Cell Biol. 130, 313-330, 1995
A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin ge

the repeat domain.

A:Reference number: Z22237; MUID:95340633

A:Accession: T42714

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1765 <PET>

A:Cross-references: EMBL:L40632; NID:g710548; PID:g710551; PIDN:AAB01605.1

A:Experimental source: strain C57BL/6J; kidney

C:Genetics:

A:Gene: Ank3

A:Map position: 10

A:Introns: 1587/1

C:Superfamily: ankyrin; ankyrin repeat homology

C:Keywords: alternative splicing

Query Match 26.8%; Score 312.5; DB 2; Length 1765;
Best Local Similarity 35.5%; Pred. No. 1.3e-18;
Matches 72; Conservative 31; Mismatches 67; Indels 33; Gaps 1;

QY 42 TALHWACSGAGTEIVEFLQLGVVNDKDDAGWSPLHIAASAGRDEIVKALLGK----- 95

DB 517 TPLHAAAREGHEDVAFLDGHGASLSITTTKGGFTPLHVAAGYKLEVASLLLOKSASPD 576

QY 96 -----GAQVNAVNGCTPLHYAASKNRHEIAVLMLEG 128

DB 577 AGKSGLTPLHVAAHYDNQKVALLLDQGSFHAAGNGYTPHIAAKNQMDIATSLLEY 636

QY 129 GANPDADHYEATAMHRAAAKGNLKMHIILLYKASTNIQDTGNTPLHLACDEERVEEA 188

DB 637 GADANAVTRQGIASVHLAAQEGHVDVMSLLSRNANVLSNKSGLTPLHLAAQEDRVNVA 696

QY 189 KLLVSOGASIYIENKEEKTPLQV 211

DB 697 EVLVNQGAVHDAQTKMGYTPLVH 719

RESULT 3

T42715

ankyrin 3, splice form 3 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000

C:Accession: T42715

R:Peterson, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo,

J. Cell Biol. 130, 313-330, 1995

A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene

the repeat domain.

A:Reference number: Z22237; MUID:95340633

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1940 <PET>

A:Cross-references: EMBL:L40632; NID:g710548; PID:g710549; PIDN:AAB01604.1

A:Experimental source: strain C57BL/6J; kidney

C:Genetics:

A:Gene: Ank3

A:Map position: 10

A:Introns: 834/1

C:Superfamily: ankyrin; ankyrin repeat homology

C:Keywords: alternative splicing

Query Match 26.8%; Score 312.5; DB 2; Length 1940;

Best Local Similarity 35.5%; Pred. No. 1.3e-18;

Matches 72; Conservative 31; Mismatches 67; Indels 33; Gaps 1;

QY 42 TALHWACSGAGTEIVEFLQLGVVNDKDDAGWSPLHIAASAGRDEIVKALLGK----- 95

DB 517 TPLHAAAREGHEDVAFLDGHGASLSITTTKGGFTPLHVAAGYKLEVASLLLOKSASPD 576

QY 96 -----GAQVNAVNGCTPLHYAASKNRHEIAVLMLEG 128

DB 577 AGKSGLTPLHVAAHYDNQKVALLLDQGSFHAAGNGYTPHIAAKNQMDIATSLLEY 636

QY 129 GANPDADHYEATAMHRAAAKGNLKMHIILLYKASTNIQDTGNTPLHLACDEERVEEA 188

DB 637 GADANAVTRQGIASVHLAAQEGHVDVMSLLSRNANVLSNKSGLTPLHLAAQEDRVNVA 696

QY 189 KLLVSOGASIYIENKEEKTPLQV 211

DB 697 EVLVNQGAVHDAQTKMGYTPLVH 719

RESULT 4

T42713

ankyrin 3, splice form 1 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000

C:Accession: T42713

R:Peterson, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo,

J. Cell Biol. 130, 313-330, 1995

A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin ge

the repeat domain.

A:Reference number: Z22237; MUID:95340633

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1943 <PET>

A:Cross-references: EMBL:L40632; NID:g710548; PID:g710550; PIDN:AAB01606.1

A:Experimental source: strain C57BL/6J; kidney

C:Genetics:

A:Gene: Ank3

A:Map position: 10

A:Introns: 855/1

C:Function:

A:Note: major kidney ankyrin

C:Superfamily: ankyrin; ankyrin repeat homology

C:Keywords: alternative splicing

Query Match 26.8%; Score 312.5; DB 2; Length 1943;

Best Local Similarity 35.5%; Pred. No. 1.3e-18;

Matches 72; Conservative 31; Mismatches 67; Indels 33; Gaps 1;

QY 42 TALHWACSGAGTEIVEFLQLGVVNDKDDAGWSPLHIAASAGRDEIVKALLGK----- 95

DB 517 TPLHAAAREGHEDVAFLDGHGASLSITTTKGGFTPLHVAAGYKLEVASLLLOKSASPD 576

QY 96 -----GAQVNAVNGCTPLHYAASKNRHEIAVLMLEG 128

DB 577 AGKSGLTPLHVAAHYDNQKVALLLDQGSFHAAGNGYTPHIAAKNQMDIATSLLEY 636

QY 129 GANPDADHYEATAMHRAAAKGNLKMHIILLYKASTNIQDTGNTPLHLACDEERVEEA 188

DB 637 GADANAVTRQGIASVHLAAQEGHVDVMSLLSRNANVLSNKSGLTPLHLAAQEDRVNVA 696

QY 189 KLLVSOGASIYIENKEEKTPLQV 211

DB 697 EVLVNQGAVHDAQTKMGYTPLVH 719

RESULT 5

T42716

ankyrin 3, splice form 4 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000

C:Accession: T42716

R:Peterson, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo,

J. Cell Biol. 130, 313-330, 1995

A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin ge

the repeat domain.

A:Reference number: Z22237; MUID:95340633

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1961 <PET>
A:Cross-references: EMBL:L40632; NID:G710548; PID:G710552; PIDN:AAB01607.1
A:Experimental source: strain C57Bl/6J; Kidney
C:Genetics:
A:Gene: Ank3
A:Map position: 10
C:Superfamily: ankryrin; ankryrin repeat homology
C:Keywords: alternative splicing

Query Match 26.8%; Score 312.5; DB 2; Length 1961;
Best Local Similarity 35.5%; Pred. No. 1.3e-18;
Matches 72; Conservative 31; Mismatches 67; Indels 33; Gaps 1;

Qy 42 TALHWACSGAGTEIVEFLQLQGVNVNDKDGASPLHIAASAGRDEIVKALLGK----- 95
Db 517 TPLHAAEGHEDVAAFLLDHGASLSITTKGFTPLHVAAYKGLKLEVASILLQKSASDA 576
Qy 96 -----GAQVNAVNGCTPLHYAASKNRHEIAVMLEGG 128
Db 577 AGKSGLTPLHVAHYDNQKVALLLDQASPHAAAKNGYTPLHIAAKKNQMDIATSLLEY 636
Qy 129 GANPPAKDHYEATAMHRAAAGNLMKMIHLLYYKASTNIQDTGNTPLHLACDERVEEA 188
Db 637 GADANAVTROGIASTVHLAAQEGHVDMSLLSRNANVLSNKSGLTPLHLAAQEDRVNVA 696
Qy 189 KLLVSOGASIVYENKEEKTPLQV 211
Db 697 EVLVNQGHAHDAQTKMGYTPLVH 719

RESULT 6
T50984
related to 26s proteasome subunit p28 [imported] - Neurospora crassa
N:Alternate names: protein B7F18.30
C:Species: Neurospora crassa
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C:Accession: T50984
R:Schulte, U.; Align, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, July 2000
A:Reference number: 225286
A:Accession: T50984
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-237 <SCH>
A:Cross-references: EMBL:AL389891; GSPDB:GN00116; NCSP:B7F18.30
A:Experimental source: BAC clone B7F18; strain OR74A
C:Genetics:
A:Gene: NCSP:B7F18.30
A:Map position: 6
A:Introns: 17/1; 25/3; 68/3

Query Match 26.6%; Score 310; DB 2; Length 237;
Best Local Similarity 34.5%; Pred. No. 1.5e-19;
Matches 78; Conservative 30; Mismatches 90; Indels 28; Gaps 4;

Qy 14 AYSKLEELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFL-QLGVPVNDKDDA 72
Db 12 ARDGKASTVESLLNANPKLAQKDDGRPLHWCASYNRKEVVELLVNQGDPDVEDDM 71
Qy 73 GWSPLHIAASA-GRDEIVKALLGGAOVNAVNGCTPLHYAASKNRHEIAVMLEGGAN 131
Db 72 GWTPEMISAKVSDSAIIDLDSRGADINQTNHOSQTAHFIAAKNMDIATSLLEYGADANAVTROG 131
Qy 132 PD-----AKDHYEATAMHRAAAGNLMKMIHLLYYKASTNIQDTGNTPLHLACDERVEE 186
Db 132 PKPASVRVKRGGYPLHRAAAGSVPMINLLQHKSPINASDNAGYTPPLHVAEGHGH 191
Qy 187 EAKLLVSOGASI-----YIENKEEKTPLQV 211
Db 192 AAVALLKAGAEATKMDGVALDLAPDKVRRFIEKAEKEGIEL 237

RESULT 7
A55575
ankryrin 3, long splice form - human
N:Alternate names: ankryrin G
C:Species: Homo sapiens (man)
C:Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 20-Sep-1999
C:Accession: A55575
R:Kordeli, E.; Lambert, S.; Bennett, V.
J. Biol. Chem. 270, 2352-2359, 1995
A:Title: Ankryrin-G. A new ankryrin gene with neural-specific isoforms localized at the
A:Reference number: A55575; MUID:95138209
A:Accession: A55575
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-4377 <KOR>
A:Cross-references: GB:U13616; NID:G608024; PIDN:AAA64834.1; PID:G608025
C:Genetics:
A:Gene: GDB:ANK3
A:Cross-references: GDB:424503; OMIM:600465
A:Map position: 10q21-10q21
C:Superfamily: unassigned ankryrin repeat proteins; ankryrin repeat homology: EGF homol
C:Keywords: alternative splicing; peripheral membrane protein

F:73-105/Domain: ankryrin repeat homology <AN01>
F:106-138/Domain: ankryrin repeat homology <AN02>
F:139-171/Domain: ankryrin repeat homology <AN03>
F:172-200/Domain: ankryrin repeat homology <AN04>
F:201-233/Domain: ankryrin repeat homology <AN05>
F:234-266/Domain: ankryrin repeat homology <AN06>
F:267-299/Domain: ankryrin repeat homology <AN07>
F:300-332/Domain: ankryrin repeat homology <AN08>
F:333-365/Domain: ankryrin repeat homology <AN09>
F:366-398/Domain: ankryrin repeat homology <AN10>
F:399-431/Domain: ankryrin repeat homology <AN11>
F:432-464/Domain: ankryrin repeat homology <AN12>
F:465-497/Domain: ankryrin repeat homology <AN13>
F:498-530/Domain: ankryrin repeat homology <AN14>
F:531-563/Domain: ankryrin repeat homology <AN15>
F:564-596/Domain: ankryrin repeat homology <AN16>
F:597-629/Domain: ankryrin repeat homology <AN17>
F:630-662/Domain: ankryrin repeat homology <AN18>
F:663-695/Domain: ankryrin repeat homology <AN19>
F:696-728/Domain: ankryrin repeat homology <AN20>
F:729-761/Domain: ankryrin repeat homology <AN21>
F:762-784/Domain: ankryrin repeat homology <AN22>
F:795-827/Domain: ankryrin repeat homology <AN23>

Query Match 26.6%; Score 310; DB 2; Length 4377;
Best Local Similarity 37.0%; Pred. No. 5.7e-18;
Matches 71; Conservative 36; Mismatches 85; Indels 0; Gaps 0;

Qy 20 BELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFLQLQGVNVNDKDDAGWSPLHI 79
Db 545 EDVAAFLLDHGASLSITTKGFTPLHVAAYKGLKLEVASILLQKSASPDAAKSGSLTPLHV 604
Qy 80 AASAGRDEIVKALLGGAOVNAVNGCTPLHYAASKNRHEIAVMLEGGANPDADKHYE 139
Db 605 AAHYDNQKVALLLDQASPHAAAKNGYTPLHIAAKKNQMDIATSLLEYGADANAVTROG 664
Qy 140 ATAMHRAAAGNLMKMIHLLYYKASTNIQDTGNTPLHLACDERVEEAKLLVSOGASIY 199
Db 665 IASVHLAAQEGHVDMSVLSLLGRNANVLSNKSGLTPLHLAAQEDRVNVAEVLVNOGHVD 724
Qy 200 IENKEEKTPLQV 211
Db 725 AQTGMGYTPLVH 736

RESULT 8
T33631
hypothetical protein F40G9.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T33631
R;Graves, T.; Sutterer, C.; Ozersky, P.
submitted to the EMBL Data Library, October 1998
A;Description: The sequence of C. elegans cosmid F40G9.
A;Reference number: 221378

A;Accession: T33631
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-636 <GRA>
A;Cross-references: EMBL:AF099919; PIDN:AAC68798.1; GSPDB:GN00021; CESP:F40G9.1
A;Experimental source: strain Bristol N2; clone F40G9
C;Genetics:
A;Gene: CESP:F40G9.1
A;Map position: 3
A;Introns: 21/3; 54/2; 90/3; 130/1; 157/3; 219/1; 260/3; 316/1; 351/3; 415/1; 438/3; 461/3

Query Match 26.2%; Score 305; DB 2; Length 636;
Best Local Similarity 29.3%; Pred. No. 1.4e-18;
Matches 86; Conservative 39; Mismatches 79; Indels 90; Gaps 6;

QY 5 VSNLWCVN-LAYSGKLEELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFLLQLG 63

Db 335 ILHKMKILHYFTKNVEARLLTRYPKLVGYTDDSGRSTHFAAVGSLPLLOFAI--- 391

QY 64 VPVNDKDDA----- 72

Db 392 --LNDPEMAHKTDVVRGMWNLTEKLENLFENFKICDFRKNYSKNTKNPHSEFFETID 449

QY 73 -----GWSPLHTAASAGDEIVKALLG-KGAQVNA 101

Db 450 FEHNSONALKKFFSLKIDIFDLNCLILPLGWTPLMTASSAGRVVVRVYLLTLPDQVVKH 509

QY 102 VNONGCTPLHYAAKKNRHETAVMLLEGGAN-PDAKHYEATAMHRAAAGNLMKMIHILLY 160

Db 510 TSNKQTSLSHYACSNHVEIKVLLEADPNIIINLPKFGATLHRAASRGNDVIRALVS 569

QY 161 Y-KASTNIQTEGNTPLHLACDEEREEAKLLVSQASIYIENKEEKTPLQVAK 213

Db 570 TGKCSLDRODQEGNTALHLACDENRGDVAILLNVNMGADMKMLNKEKQTPLEMLK 623

RESULT 9
S37431
N;Alternate names: ankyrin B, 440K splice form; ankyrin-B; brain ankyrin; non-erythroid
N;Contains: ankyrin 2, short form
C;Species: Homo sapiens (man)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 13-Aug-1999
C;Accession: S37431; A39643; B39643; A40334; A49462; S14533; S14569
R;Chan, W.

submitted to the EMBL Data Library, September 1993
A;Reference number: S37431
A;Accession: S37431
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-3924 <CHA>

A;Cross-references: EMBL:Z26634; NID:g406287; PIDN:CAA81387.1; PID:g406288

R;Otto, E.; Kunimoto, M.; McLaughlin, T.; Bennett, V.

J. Cell Biol. 114, 241-253, 1991

A;Title: Isolation and characterization of cDNAs encoding human brain ankyrins reveal a

A;Reference number: A39643; MUID:91302466

A;Accession: A39643

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-2077 <OTI>

A;Cross-references: GB:X56957

A;Accession: B39643

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1443, 3585-3924 <OTT>

A;Cross-references: EMBL:X56958

R;Tse, W.T.; Menninger, J.C.; Yang-Feng, T.L.; Francke, U.; Sahr, K.E.; Lux, S.E.; Wa
genomics 10, 858-866, 1991

A;Title: Isolation and chromosomal localization of a novel nonerythroid ankyrin gene.

A;Reference number: A40334; MUID:92009921

A;Accession: A40334

A;Molecule type: DNA

A;Residues: 463-474, 'PE', 477-495 <TSE>

A;Cross-references: GB:M37123; NID:g178647; PIDN:AAA62828.1; PID:g178648

R;Chan, W.; Kordeli, E.; Bennett, V.

J. Cell Biol. 123, 1463-1473, 1993

A;Title: 440-kD ankyrinB: structure of the major developmentally regulated domain and

A;Reference number: A49462; MUID:94075409

A;Accession: A49462

A;Status: preliminary; nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-3924 <RES>

A;Cross-references: EMBL:Z26634; NID:g406287; PIDN:CAA81387.1; PID:g406288

C;Genetics:

A;Gene: GDB:ANK2

A;Cross-references: GDB:127607; OMIM:106410

A;Map position: 4q25-4q27

C;Superfamily: ankyrin; ankyrin repeat homology

C;Keywords: alternative splicing

F;2-3924/Product: ankyrin 2, long form #status predicted <MAT>

F;2-1443, 3585-3924/Product: ankyrin 2, short form #status predicted <MA2>

F;63-95/Domain: ankyrin repeat homology <AN01>

F;96-128/Domain: ankyrin repeat homology <AN02>

F;129-161/Domain: ankyrin repeat homology <AN03>

F;162-190/Domain: ankyrin repeat homology <AN04>

F;191-223/Domain: ankyrin repeat homology <AN05>

F;232-264/Domain: ankyrin repeat homology <AN06>

F;265-297/Domain: ankyrin repeat homology <AN07>

F;298-330/Domain: ankyrin repeat homology <AN08>

F;331-363/Domain: ankyrin repeat homology <AN09>

F;364-396/Domain: ankyrin repeat homology <AN10>

F;397-429/Domain: ankyrin repeat homology <AN11>

F;430-462/Domain: ankyrin repeat homology <AN12>

F;463-495/Domain: ankyrin repeat homology <AN13>

F;496-528/Domain: ankyrin repeat homology <AN14>

F;529-561/Domain: ankyrin repeat homology <AN15>

F;562-594/Domain: ankyrin repeat homology <AN16>

F;595-627/Domain: ankyrin repeat homology <AN17>

F;628-660/Domain: ankyrin repeat homology <AN18>

F;661-693/Domain: ankyrin repeat homology <AN19>

F;694-726/Domain: ankyrin repeat homology <AN20>

F;727-759/Domain: ankyrin repeat homology <AN21>

F;760-792/Domain: ankyrin repeat homology <AN22>

F;793-825/Domain: ankyrin repeat homology <AN23>

Query Match 25.3%; Score 294; DB 2; Length 3924;
Best Local Similarity 31.9%; Pred. No. 1.2e-16;
Matches 74; Conservative 40; Mismatches 84; Indels 34; Gaps 2;

QY 14 AYSGKLEELKESILADKSLATRTDQDSRTALHWACSGAGTEIVEFLLQ----- 61

Db 472 ARAGQVEVVR-CLLRNGALVDARAREEQTPHTASRLGKTEIVQLLQHHMAHPDAATTNG 530

QY 62 -----LQVPVNDKDDAGWSPLHTAASAGRDEIVKALLGKAQVN 100

Db 531 YTPLHTSAREGGVDVASVLLLEAGAAHSLATKGTTPHLVAAKYGSLDVAKLLQRRRAAD 590

QY 101 AVNONGCTPLHYAASKNRHEIAVMLLEGGANPDADKHYEATAMHRAAAGNLMKMIHILLY 160

Db 591 SACKNGLTPLHYAAHYDNQKVALLLLEKASPHATKNGYTPHTIAAKKNQMQIASTLLN 650

QY 161 YKASTNIQTEGNTPLHLACDEEREEAKLLVSQASIYIENKEEKTPLQOVA 212

Db 651 YGAETNIVTKQGVTPHLHLASQEGHTDMVTLLLDKGANIHMSTKSGLTSLHLA 702

RESULT 10
S57697

hypothetical protein YGR232w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein G8564
C:Species: Saccharomyces cerevisiae
C:Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 23-Mar-2001
C:Accession: S57697; S64556; S63913
R:van der Aart, Q.J.M.; Kleine, K.; Steensma, H.Y.
A:Submitted to the EMBL Data Library, June 1995
A:Description: Sequence analysis of the 43 kb CRM1-YLM9-PET54-SMI1-PHO81-YHB4-PFK1 region
A:Reference number: S57680
A:Accession: S57697
A:Molecule type: DNA
A:Residues: 1-228 <VAN>
A:Cross-references: EMBL:X87941; NID:G886908; PIDN:CAA61182.1; PID:G886926
A:Experimental source: strain S288C
R:van der Aart, Q.J.M.; Steensma, H.Y.
A:Submitted to the Protein Sequence Database, May 1996.
A:Reference number: S64541
A:Accession: S64556
A:Molecule type: DNA
A:Residues: 1-228 <VAN>
A:Cross-references: EMBL:Z73017; NID:g1323418; PIDN:CAA97260.1; PID:e243668; PID:g132341
A:Experimental source: strain S288C
R:van der Aart, Q.J.M.; Kleine, K.; Steensma, H.Y.
A:Title: Sequence analysis of the 43 kb CRM1-YLM9-PET54-DIE2-SMI1-PHO81-YHB4-PFK1 region
A:Reference number: S63896; MUID:96267763
A:Accession: S63913
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-228 <VAN>
A:Cross-references: EMBL:X87941; NID:G886908; PIDN:CAA61182.1; PID:G886926
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995
C:Genetics:
A:Map position: 7R
A:Note: YGR232w
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
F:71-103/Domain: ankyrin repeat homology <AN1>

Query Match 25.0%; Score 291; DB 2; Length 228;
Best Local Similarity 35.7%; Pred. No. 6.4e-18;
Matches 75; Conservative 36; Mismatches 91; Indels 8; Gaps 4;
Qy 5 VSNLMVCNLAISGKLEELKESLADKSLATRTDQDSRTALHWACSGAGTEHTEIFLLQGV 64
Db 1 MSNYPHQAQCHNEFFKVOELHSPSLQLQDQDGRPLHWSVSFQAHEITSFLLSKME 60
Qy 65 PVN---DRDDAGWSPLHIAASAGRDEIVKALGKGAQ--VNAVNGCGCTPLHYAASKNRH 119
Db 61 NVNLDYDPSDGSWTFPHIACSVGNLEVKSLYDRPLKPDNKNITNQGVTCLEHLAVGKKWF 120
Qy 120 ETAVMLLEGGANPDADHYEATAMHRAAKGNLKMIIHLLYYKASTNIQDTTEGNTPLHL 178
Db 121 EVSQPLIENGASVRILKDFNQIPLHRAASVSGSLKILIELCLGLGKSAVNQDQGWTPLFH 180
Qy 179 ACDEERVEEAKLLVSQASIV--IENKEEK 206
Db 181 ALAEGHGAVALLVKRYGAEDYLDVNKGAK 210

RESULT 11
B35049
ankyrin 1, erythrocyte splice form 3 - human
N:Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
N:Contains: ankyrin 2.2, erythrocyte
C:Species: Homo sapiens (man)
C:Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 10-Jul-1998
C:Accession: B35049
R:Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.;
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A:Title: cDNA sequence for human erythrocyte ankyrin.
A:Reference number: A35049; MUID:90175370
A:Accession: B35049

Query Match 24.9%; Score 290; DB 2; Length 1856;
Best Local Similarity 34.3%; Pred. No. 1e-16;
Matches 79; Conservative 38; Mismatches 105; Indels 8; Gaps 3;
Qy 5 VSNLMV---CNLAISGKLEELKESLADKSLATRTDQDSRTALHWACSGAGTEHTEIFLLQ 61
Db 432 VSNVAVETPLHMAARAGHTEVAKYLLQNKAKYNAKDDQTPHCAARIGHTNMVKLLLE 491
Qy 62 LGVPVNDKDDAGWSPLHIAASAGRDEIVKALGKGAQVNAVNGCGCTPLHYAASKNRHEI 121
Db 492 NNANPLATTATTPHIAAREGHVETVLALLEKASOACMTKGTPLHVAAKYGVKVRV 551
Qy 122 AVMLLEGGANPDADHYEATAMHRAAKGNLKMIIHLLYYKASTNIQDTTEGNTPLHLACD 181
Db 552 AELLERDAHPNAACNGITPLHVAHHNNLDIVKLLLPGRGSPHSPANNVYTPHIAAK 611
Qy 182 ERVEEAKLLVSQASIVYENKEETPLQVA--KGG---LGUILKRWVEG 226
Db 612 QNQVEVARSLQYGGSAANAESVQGVTPHLAAQAGHAEMVALLSKQANG 661

RESULT 12
A35049
ankyrin 1, erythrocyte splice form 2 - human
N:Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
N:Contains: ankyrin 2.2, erythrocyte
C:Species: Homo sapiens (man)
C:Date: 27-Jul-1990 #sequence_revision 01-Oct-1992 #text_change 04-Sep-1998
C:Accession: A35049
R:Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.;
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A:Title: cDNA sequence for human erythrocyte ankyrin.
A:Reference number: A35049; MUID:90175370
A:Accession: A35049
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1880 <LAM>

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1856 <LAM>
C:Genetics:
A:Gene: GDB:ANK1; ANK
A:Cross-references: GDB:118737; OMIM:182900
A:Map position: 8p11.2-8p11.2
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing
F:2-1856/Product: ankyrin 1, erythrocyte form 3 #status predicted <MA2>
F:2-1513,1676-1856/Product: ankyrin 2.2, erythrocyte
F:44-76/Domain: ankyrin repeat homology <AN01>
F:77-109/Domain: ankyrin repeat homology <AN02>
F:110-142/Domain: ankyrin repeat homology <AN03>
F:143-171/Domain: ankyrin repeat homology <AN04>
F:172-204/Domain: ankyrin repeat homology <AN05>
F:205-237/Domain: ankyrin repeat homology <AN06>
F:238-270/Domain: ankyrin repeat homology <AN07>
F:271-303/Domain: ankyrin repeat homology <AN08>
F:304-336/Domain: ankyrin repeat homology <AN09>
F:337-369/Domain: ankyrin repeat homology <AN10>
F:370-402/Domain: ankyrin repeat homology <AN11>
F:403-435/Domain: ankyrin repeat homology <AN12>
F:436-468/Domain: ankyrin repeat homology <AN13>
F:469-501/Domain: ankyrin repeat homology <AN14>
F:502-534/Domain: ankyrin repeat homology <AN15>
F:535-567/Domain: ankyrin repeat homology <AN16>
F:568-600/Domain: ankyrin repeat homology <AN17>
F:601-633/Domain: ankyrin repeat homology <AN18>
F:634-666/Domain: ankyrin repeat homology <AN19>
F:667-699/Domain: ankyrin repeat homology <AN20>
F:700-732/Domain: ankyrin repeat homology <AN21>
F:733-765/Domain: ankyrin repeat homology <AN22>
F:766-798/Domain: ankyrin repeat homology <AN23>

A;Cross-references: GB:M28880

C;Genetics:

A;Gene: GDB:ANK1; ANK

A;Cross-references: GDB:118737; OMIM:182900

A;Map Position: 8p11.2-8p11.2

C;Superfamily: ankyrin; ankyrin repeat homology

C;Keywords: alternative splicing; cytoskeleton

F;2-1880/Product: ankyrin 1, erythrocyte form 2 #status predicted <MAT>

F;2-1513,1676-1880/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>

F;44-76/Domain: ankyrin repeat homology <AN01>

F;77-109/Domain: ankyrin repeat homology <AN02>

F;110-142/Domain: ankyrin repeat homology <AN03>

F;143-171/Domain: ankyrin repeat homology <AN04>

F;172-204/Domain: ankyrin repeat homology <AN05>

F;205-237/Domain: ankyrin repeat homology <AN06>

F;238-270/Domain: ankyrin repeat homology <AN07>

F;271-303/Domain: ankyrin repeat homology <AN08>

F;304-336/Domain: ankyrin repeat homology <AN09>

F;337-369/Domain: ankyrin repeat homology <AN10>

F;370-402/Domain: ankyrin repeat homology <AN11>

F;403-435/Domain: ankyrin repeat homology <AN12>

F;436-468/Domain: ankyrin repeat homology <AN13>

F;469-501/Domain: ankyrin repeat homology <AN14>

F;502-534/Domain: ankyrin repeat homology <AN15>

F;535-567/Domain: ankyrin repeat homology <AN16>

F;568-600/Domain: ankyrin repeat homology <AN17>

F;601-633/Domain: ankyrin repeat homology <AN18>

F;634-666/Domain: ankyrin repeat homology <AN19>

F;667-699/Domain: ankyrin repeat homology <AN20>

F;700-732/Domain: ankyrin repeat homology <AN21>

F;733-765/Domain: ankyrin repeat homology <AN22>

F;766-798/Domain: ankyrin repeat homology <AN23>

Query Match 24.9%; Score 290; DB 2; Length 1880;

Best Local Similarity 34.3%; Pred. No. 1.1e-16; 3;

Matches 79; Conservative 38; Mismatches 105; Indels 8; Gaps 3;

QY 5 VSNLMV---CNLAYSGKLEELKESILADKSLATRTDSDSTALHWACSGAGTEIVEFLQ 61

Db 432 VSNKVTETPLHMAARAGHTVAKYLLQNKAKVNAKDDQTPHCAARIGHTNMVKILLE 491

QY 62 LGVPVNDKDDAGWSPLHIAASAGRDEIVKALLGKGAOVNAVNGCTPLHYAASKNRHEI 121

Db 492 NNANPNLATTAGHTPLHMAARAGHTVAKYLLQNKAKVNAKDDQTPHCAARIGHTNMVKILLE 491

QY 122 AVMLLEGGANPDADKHVEATAMHRAAKGNLKMHIILLYKASTNIQDTGNTPLHLACD 181

Db 552 AELLERDAHPNAAGKNGLTPLHVAVHNNLDIVKLLPRGGSPHSPAWNNGYTPPLHIAAK 611

QY 182 EERVEEAKLLVSOGASTIYENKEEKTPLQVA--KGG--LGLILKRMVEG 226

Db 612 QNQEVARSLQYGGSANAESVQGTPLHMAARAGHTVAKYLLQNKAKVNAKDDQTPHCAARIGHTNMVKILLE 491

QY 182 EERVEEAKLLVSOGASTIYENKEEKTPLQVA--KGG--LGLILKRMVEG 226

Db 612 QNQEVARSLQYGGSANAESVQGTPLHMAARAGHTVAKYLLQNKAKVNAKDDQTPHCAARIGHTNMVKILLE 491

QY 182 EERVEEAKLLVSOGASTIYENKEEKTPLQVA--KGG--LGLILKRMVEG 226

Db 612 QNQEVARSLQYGGSANAESVQGTPLHMAARAGHTVAKYLLQNKAKVNAKDDQTPHCAARIGHTNMVKILLE 491

QY 182 EERVEEAKLLVSOGASTIYENKEEKTPLQVA--KGG--LGLILKRMVEG 226

Db 612 QNQEVARSLQYGGSANAESVQGTPLHMAARAGHTVAKYLLQNKAKVNAKDDQTPHCAARIGHTNMVKILLE 491

QY 182 EERVEEAKLLVSOGASTIYENKEEKTPLQVA--KGG--LGLILKRMVEG 226

Db 612 QNQEVARSLQYGGSANAESVQGTPLHMAARAGHTVAKYLLQNKAKVNAKDDQTPHCAARIGHTNMVKILLE 491

QY 182 EERVEEAKLLVSOGASTIYENKEEKTPLQVA--KGG--LGLILKRMVEG 226

Db 612 QNQEVARSLQYGGSANAESVQGTPLHMAARAGHTVAKYLLQNKAKVNAKDDQTPHCAARIGHTNMVKILLE 491

QY 182 EERVEEAKLLVSOGASTIYENKEEKTPLQVA--KGG--LGLILKRMVEG 226

Db 612 QNQEVARSLQYGGSANAESVQGTPLHMAARAGHTVAKYLLQNKAKVNAKDDQTPHCAARIGHTNMVKILLE 491

QY 182 EERVEEAKLLVSOGASTIYENKEEKTPLQVA--KGG--LGLILKRMVEG 226

Db 612 QNQEVARSLQYGGSANAESVQGTPLHMAARAGHTVAKYLLQNKAKVNAKDDQTPHCAARIGHTNMVKILLE 491

QY 182 EERVEEAKLLVSOGASTIYENKEEKTPLQVA--KGG--LGLILKRMVEG 226

Db 612 QNQEVARSLQYGGSANAESVQGTPLHMAARAGHTVAKYLLQNKAKVNAKDDQTPHCAARIGHTNMVKILLE 491

QY 182 EERVEEAKLLVSOGASTIYENKEEKTPLQVA--KGG--LGLILKRMVEG 226

Db 612 QNQEVARSLQYGGSANAESVQGTPLHMAARAGHTVAKYLLQNKAKVNAKDDQTPHCAARIGHTNMVKILLE 491

A;Residues: 2-7,'X',9-17,'X',19-20,'T',22-30;733-749,'A',751-753;828-833,'X',835-855,

X',1367;1383-1427;1601-1630;1686-1698,'D',1700;1763-1772 <LOX>

A;Note: 845-Arg and 1392-Thr were also found

R;Hermann, J.; Barel, M.; Frade, R.

Biochem. Biophys. Res. Commun. 204, 453-460, 1994

A;Title: Human erythrocyte ankyrin, a cytoskeleton component, generates the p57 membr

A;Reference number: PC2220; MUID:95071348

A;Accession: PC2220

A;Molecule type: protein

A;Residues: 910-929 <HER>

R;Davis, L.H.; Bennett, V.

J. Biol. Chem. 265, 10589-10596, 1990

A;Title: Mapping the binding sites of human erythrocyte ankyrin for the anion exchang

A;Reference number: A35443; MUID:90285190

A;Accession: A35443

A;Molecule type: protein

A;Residues: 'X',5,'X',7-12,403-417,'X',419-422,'H',424,'LQ';797-800,'L',802-814;862-8

C;Genetics:

A;Gene: GDB:ANK1; ANK

A;Cross-references: GDB:118737; OMIM:182900

A;Map position: 8p11.2-8p11.2

C;Superfamily: ankyrin; ankyrin repeat homology

C;Keywords: alternative splicing; phosphoprotein

F;2-1881/Product: ankyrin 1, erythrocyte form 1 #status predicted <MAT1>

F;2-1512,1675-1881/Product: ankyrin 2.2, erythrocyte #status predicted <MAT2>

F;2-827/Domain: 89K #status predicted <DOM1>

F;2-827/Region: anion exchange protein binding

F;44-76/Domain: ankyrin repeat homology <AN01>

F;77-109/Domain: ankyrin repeat homology <AN02>

F;110-142/Domain: ankyrin repeat homology <AN03>

F;143-171/Domain: ankyrin repeat homology <AN04>

F;172-204/Domain: ankyrin repeat homology <AN05>

F;205-237/Domain: ankyrin repeat homology <AN06>

F;238-270/Domain: ankyrin repeat homology <AN07>

F;271-303/Domain: ankyrin repeat homology <AN08>

F;304-336/Domain: ankyrin repeat homology <AN09>

F;337-369/Domain: ankyrin repeat homology <AN10>

F;370-402/Domain: ankyrin repeat homology <AN11>

F;403-435/Domain: ankyrin repeat homology <AN12>

F;436-468/Domain: ankyrin repeat homology <AN13>

F;469-501/Domain: ankyrin repeat homology <AN14>

F;502-534/Domain: ankyrin repeat homology <AN15>

F;535-567/Domain: ankyrin repeat homology <AN16>

F;568-600/Domain: ankyrin repeat homology <AN17>

F;601-633/Domain: ankyrin repeat homology <AN18>

F;634-666/Domain: ankyrin repeat homology <AN19>

F;667-699/Domain: ankyrin repeat homology <AN20>

F;700-732/Domain: ankyrin repeat homology <AN21>

F;733-765/Domain: ankyrin repeat homology <AN22>

F;766-798/Domain: ankyrin repeat homology <AN23>

F;828-1382/Domain: 62K #status predicted <DOM2>

F;828-1382/Region: spectrin binding

F;1383-1881/Domain: 55K #status predicted <DOM3>

Db	612	QNVQVEARSLQYGGSSANAESVQGVTPHLHAAQEGHAEMVALLSKQANG	661
RESULT	14		
A:57282			
<p>ankyrin-related protein unc-44 - Caenorhabditis elegans (fragment)</p> <p>C:Species: Caenorhabditis elegans</p> <p>C>Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 20-Sep-1999</p> <p>C:Accession: A57282</p> <p>R:Otsuka, A.J.; Franco, R.; Yang, B.; Shim, K.H.; Tang, L.Z.; Zhang, Y.Y.; Boontrakulpo</p> <p>J. Cell Biol. 129, 1081-1092, 1995</p> <p>A:Title: An ankyrin-related gene (unc-44) is necessary for proper axonal guidance in Cae</p> <p>A:Reference number: A57282; MUID:95263663</p> <p>A:Accession: A57282</p> <p>A>Status: preliminary</p> <p>A:Molecule type: DNA</p> <p>A:Residues: 1-1786 <OTS></p> <p>A:Cross-references: GB:U21734; NID:g790607; PIDN:AAA85854.1; PID:g790608</p> <p>C:Genetics:</p> <p>A:Gene: unc-44</p> <p>A:Introns: 97/3; 262/2; 535/1; 749/1; 786/3; 883/3; 1001/2; 1007/2; 1051/3; 1666/1; 1725</p> <p>C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology</p> <p>F:164-162/Domain: ankyrin repeat homology <AN04></p> <p>F:358-390/Domain: ankyrin repeat homology <AN1></p> <p>F:391-423/Domain: ankyrin repeat homology <AN11></p>			
Query Match	24.8%	Score 288.5; DB 2; Length 1786;	
Best Local Similarity	36.0%	Pred. No. 1.3e-16;	
Matches	76; Conservative	37; Mismatches 89; Indels	9; Gaps 4;
Qy	20	BEKESIL---ADKSLATRTDQDSRTALHWACSGAGTEIVEFLLOLGVVPVNDKDDAGWSP	76
Db	537	EEVAGILDHNDKTLTK---KGFTPLHLASKYGNLEVRLLLERGTPVDIEGKNQVTP	593
Qy	77	LHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHEIAVMLEGGANPDARD	136
Db	594	LHVAAHYNNDKVAMLENGASAKAAKNGYTPPLHIAAKKNQMEIASTLLQFKADPNNAKS	653
Qy	137	HYEATAMHRAAAKGNLKHILLYKASTNQDTGNTPLHLACDEERVEEAKLLVSOGA	196
Db	654	RAGFTPLHLSAQEGHKEISGLLIENGSDVGAKANNGLTAMHLCAQEDHVPVQAIIYNNGA	713
Qy	197	SIYIENKEEKTPLQVA---KGGGLGLILKRWVE	225
Db	714	EINSKTNAGYTPPLHVACHFGQLNMV-KFLVE	743
<p>Search completed: August 13, 2001, 07:46:40</p> <p>Job time: 137 sec</p>			
RESULT	15		
T15346			
<p>elegans ankyrin-related unc-44 - Caenorhabditis elegans</p> <p>C:Species: Caenorhabditis elegans</p> <p>C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jan-2000</p> <p>C:Accession: T15346</p> <p>R:Gattung, S.</p> <p>submitted to the EMBL Data Library, February 1996</p> <p>A:Description: The sequence of C. elegans cosmid B0350.</p> <p>A:Reference number: Z18332</p> <p>A:Accession: T15346</p> <p>A>Status: preliminary</p> <p>A:Molecule type: DNA</p> <p>A:Residues: 1-1815 <GAT></p> <p>A:Cross-references: EMBL:U50071; NID:g1208871; PID:g1208875; PIDN:AAA93445.1; CESP:unc-4</p> <p>C:Genetics:</p> <p>A:Gene: CESP:unc-44</p> <p>A:Introns: 97/3; 262/2; 535/1; 749/1; 786/3; 883/3; 1001/2; 1007/2; 1051/3; 1666/1; 1725</p> <p>C:Superfamily: ankyrin; ankyrin repeat homology</p> <p>F:358-390/Domain: ankyrin repeat homology <ANR></p>			
Query Match	24.8%	Score 288.5; DB 2; Length 1815;	
Best Local Similarity	36.0%	Pred. No. 1.4e-16;	
Matches	76; Conservative	37; Mismatches 89; Indels	9; Gaps 4;

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2001, 07:44:23 ; Search time 19.4 seconds
(without alignments)
239.867 Million cell updates/sec

Title: US-09-509-775-2
Perfect score: 1164
Sequence: 1 MEGCVSNLMVCNLAYSGKLE.....TPLQVAKGGLILKRMVEG 226

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA.*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pap.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pap.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pap.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pap.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pap.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	310	26.6	1088	4	US-09-082-059-2
2	294	25.3	1839	2	US-09-172-977-4
3	287	24.7	352	3	US-09-065-474-139
4	287	24.7	1745	2	US-09-031-485-33
5	287	24.7	1745	2	US-08-847-429A-33
6	287	24.7	1745	3	US-09-065-474-33
7	286	24.6	843	2	US-09-172-977-3
8	271	23.3	302	2	US-09-031-485-38
9	271	23.3	302	3	US-08-847-429A-38
10	271	23.3	302	3	US-09-065-474-38
11	262	22.5	303	2	US-09-031-485-23
12	262	22.5	303	2	US-08-847-429A-23
13	262	22.5	303	3	US-09-065-474-23
14	258.5	22.2	787	4	US-09-188-930-334
15	247	21.2	1423	4	US-08-810-712-10
16	242.5	20.8	752	1	US-08-281-193-2
17	242.5	20.8	752	1	US-08-422-106-2
18	242.5	20.8	752	2	US-08-735-716-2
19	242.5	20.8	752	2	US-08-555-568B-2
20	242.5	20.8	752	5	PCT-US95-08069-2
21	241	20.7	348	2	US-09-031-485-28
22	241	20.7	348	2	US-08-847-429A-28
23	241	20.7	348	3	US-09-065-474-28
24	237	20.4	387	2	US-08-484-575A-7
25	237	20.4	387	3	US-08-477-459-7
26	237	20.4	387	3	US-08-479-869-7
27	237	20.4	387	4	US-08-486-414-7

28	237	20.4	387	5	PCT-US94-01826A-7
29	237	20.4	387	5	PCT-US94-02252A-7
30	227.5	19.5	394	2	US-08-555-568B-17
31	227.5	19.5	687	2	US-08-555-568B-21
32	227.5	19.5	688	2	US-08-555-568B-23
33	226.5	19.5	191	2	US-09-031-485-20
34	226.5	19.5	191	2	US-08-847-429A-20
35	226.5	19.5	191	3	US-09-065-474-20
36	224.5	19.3	899	1	US-08-365-689-2
37	224.5	19.3	899	1	US-08-145-138A-2
38	224.5	19.3	933	1	US-07-747-781-2
39	224.5	19.3	933	5	PCT-US92-06888-2
40	212.5	18.3	657	1	US-08-264-534-34
41	212.5	18.3	657	1	US-08-083-590A-13
42	212.5	18.3	657	1	US-08-465-500-34
43	212.5	18.3	657	2	US-08-346-128-34
44	212.5	18.3	657	3	US-08-532-384-13
45	212.5	18.3	657	3	US-08-893-828-34

ALIGNMENTS

RESULT 1
US-09-082-059-2
; Sequence 2, Application US/09082059A
; Patent No. 6225086
; GENERAL INFORMATION:
; APPLICANT: Morrow, Jon S.
; APPLICANT: Devorajan, Prasad
; TITLE OF INVENTION: NO. 6225086el Ankyrin Proteins and a Method for Their Identifi
; FILE REFERENCE: 44574-5002-US
; CURRENT APPLICATION NUMBER: US/09/082,059A
; CURRENT FILING DATE: 1998-05-21
; EARLIER APPLICATION NUMBER: 60/047356
; EARLIER FILING DATE: 1997-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1088
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-082-059-2

Query Match 26.6%; Score 310; DB 4; Length 1088;
Best Local Similarity 37.0%; Pred. No. 1.9e-26;
Matches 71; Conservative 36; Mismatches 85; Indels 0; Gaps 0;
QY 20 RELKESILADKSLATRTDQDSRTALHWACSGHTEIVEFLQLGLGVFVNDKDDAGWSPLHI 79
Db 166 EDVAFLDHGASLSITTKKGTPLHVAKYCKLEVANLLQKSPDAGKSGLTPLHV 225
QY 80 AASAGRDEITVALLGKGAOVNAVNGCTPLHYAASKNRHEIAVMLLEGANPDARDHYE 139
Db 226 AAHYDNQKVALLLDQASPHAAKNGYTPHIAAKNQMDIATTLLEYGADANAVTROG 285
QY 140 ATAMHRAAKGNKMIHILLYKASTNTQDTGCTNPLHLACDEERVEEAKLLVSGASTY 199
Db 286 TASVHLAAQEGHDMVSLLLGRNANVNLNSKSLTPLHLAAQEDRYNVAEVLVNOGAHYD 345
QY 200 IENKEETPLQV 211
Db 346 AQTRMGYTPLVH 357

RESULT 2
US-09-172-977-4
; Sequence 4, Application US/09172977
; Patent No. 5989863
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.

```

; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
; FILE REFERENCE: PF-0615 US
; CURRENT APPLICATION NUMBER: US/09/172,977
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 1839
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: g29491
US-09-172-977-4

Query Match 25.3%; Score 294; DB 2; Length 1839;
Best Local Similarity 31.9%; Pred. No. 2.9e-24;
Matches 74; Conservative 40; Mismatches 84; Indels 34; Gaps 2;

QY 14 AVSGKLEELKESILAKSLATRTDQSTALHWACSAGHTEIVEFLQ----- 61
DB 472 ARAGQVEVVR-CILRRNGALVDARAEQETPLHIASRLGKTEIVQLLQHMHPDAATNG 530
QY 62 -----LGVPNDKDDAGWSPHIAASAGRDEIVKALLGKGQVN 100
DB 531 YTPLHSAREGQVDVASVLEGAHSLATKGGFTPLHVAARYGSLDVAKLLQRRRAAD 590
QY 101 AVNQNGCTPLHYAASKNRHIEAVMLLEGGANPDAKDHYEATAMHRAAAKGNLKMHIILY 160
DB 591 SACKNGLTPLHVAAHYDNQKVALLLLEKGGASHATAKNGYTPLHIAAKKNQMIASITLLN 650
QY 161 YKASTNIQTEGNTPLHLACDERVEEAKLLYSQASIASIYENKEEKTPLQVA 212
DB 651 YGAETNIVTKQGVTPPLHLSAQSGHGTDMVTLLLDKGANHMTSKSLTSLHLA 702

RESULT 3
US-09-065-474-139
; Sequence 139, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505

; INFORMATION FOR SEQ ID NO: 139:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-09-065-474-139

Query Match 24.7%; Score 287; DB 3; Length 352;
Best Local Similarity 36.5%; Pred. No. 1.4e-24;
Matches 73; Conservative 33; Mismatches 84; Indels 10; Gaps 2;

QY 20 EELKESILA-----DKSLATRTDQSTALHWACSAGHTEIVEFLQLGVPNVDKDDA 72
DB 86 KEGQEEVAAAILMDHGTDKLTLLTK---KGFTPLHLAAKYGNLVPVAKSLUERPTVDIEGKN 142
QY 73 GNSPLHIAAASAGRDEIVKALLGKGQVNAVNONCGTPTPLHYAASKNRHIEAVMLLEGGANP 132
DB 143 QVTPLHVAAHYNDKVALLLLENGASAHAAKNGYTPHIAAKKNQMIASITLLHYKANA 202
QY 133 DAKDHYEATAMHRAAAKGNLKMHIILYKASTNTQDTEGNTPLHLACDERVEEAKLLV 192
DB 203 NAEKAGFTPLHAAQEGHREMAALLIENGAKVGAQARNGLTPMHLCAQEDRVSAEELV 262
QY 193 SQASIASIYENKEEKTPLQVA 212
DB 263 KENAAIDPTKAGYTPHVA 282

RESULT 4
US-09-031-485-33
; Sequence 33, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; FILING DATE: US/09/031,485
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429
; FILING DATE: 24-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```


;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/847.429A
;; FILING DATE: 24-APR-1997
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Verser, Carol Talkington
;; REGISTRATION NUMBER: 37,459
;; REFERENCE/DOCKET NUMBER: HW-5
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 970/493-7272
;; TELEFAX: 970/484-9505
;; INFORMATION FOR SEQ ID NO: 38:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 302 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-847-429A-38

Query Match 23.3%; Score 271; DB 2; Length 302;
Best Local Similarity 35.4%; Pred. No. 7.4e-23;
Matches 67; Conservative 34; Mismatches 88; Indels 0; Gaps 0;

Qy 24 ESILADKSLATRTDQDSRTALHWACSAGHTEIVEFLQLGVVNDKDDAGWSPLHIAASA 83
Db 57 ELLLYHAAIEATTESGLSPLHVAAFMGAINIIVYLLOOGANANVATVRGETPLHLAARA 116

Qy 84 GRDEIVKALLGKAQVNAVNGCTPLHYAASKNRHETAVMLLEGGANPDADKHYEATAM 143
Db 117 NOTDIVRVLRNGAQVDAARELQTPPLHIASRLGNTDIVILLQADASPNAATRDLYTL 176

Qy 144 HRAAKGNLKMTHILLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVSOGASIYIENK 203
Db 177 HIAAKEGEEVAAIILDHGSKTLTKGFTPLHLAAKYGNLPVAKLLERGTGLVDIEGK 236

Qy 204 EEKTPLOVA 212
Db 237 NQVTPPLHVA 245

RESULT 10
US-09-065-474-38
; Sequence 38, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065.474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1

;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 970/493-7272
;; TELEFAX: 970/484-9505
;; INFORMATION FOR SEQ ID NO: 38:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 302 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-09-065-474-38

Query Match 23.3%; Score 271; DB 3; Length 302;
Best Local Similarity 35.4%; Pred. No. 7.4e-23;
Matches 67; Conservative 34; Mismatches 88; Indels 0; Gaps 0;

Qy 24 ESILADKSLATRTDQDSRTALHWACSAGHTEIVEFLQLGVVNDKDDAGWSPLHIAASA 83
Db 57 ELLLYHAAIEATTESGLSPLHVAAFMGAINIIVYLLOOGANANVATVRGETPLHLAARA 116

Qy 84 GRDEIVKALLGKAQVNAVNGCTPLHYAASKNRHETAVMLLEGGANPDADKHYEATAM 143
Db 117 NOTDIVRVLRNGAQVDAARELQTPPLHIASRLGNTDIVILLQADASPNAATRDLYTL 176

Qy 144 HRAAKGNLKMTHILLYYKASTNIQDTGNTPLHLACDEERVEEAKLLVSOGASIYIENK 203
Db 177 HIAAKEGEEVAAIILDHGSKTLTKGFTPLHLAAKYGNLPVAKLLERGTGLVDIEGK 236

Qy 204 EEKTPLOVA 212
Db 237 NQVTPPLHVA 245

RESULT 11
US-09-031-485-23
; Sequence 23, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031.485
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/847.429
; FILING DATE: 24-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 303 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-031-485-23

Query Match 22.5%; Score 262; DB 2; Length 303;
Best Local Similarity 33.9%; Pred. No. 7.9e-22;
Matches 74; Conservative 35; Mismatches 97; Indels 12; Gaps 4;
QY 14 AYSKLEELKESILADKSLATRTDQDSR-----TALHWASAGHTEIVEFLQLQGVVND 68
Db 16 AHCGHVRVAK--LLLDNRN---ADPNARALNGFTPLHIACKKNRIKIVELLKYHAAIEA 69
QY 69 KDDAGWSPLHIAASAGRDEIVKALLGKGAQVNAVQNGCTPLHYAASKNRHEIAVMLEGG 128
Db 70 TTESGLSPLHVAAPMGAINIVYLLQQGANADVATVGTETPLHLAARANQTDIVRVLRN 129
QY 129 GANPDADKHYEATAMHRAAAGNLMKMIHILLYKASTNIQDTEGNTPLHLACDEERVEEA 188
Db 130 GAQVDAARAARELQPLHIASRLGNTDIVILLQANASPNAAATRDLYTPLHIAAKEGQEEVA 189
QY 189 KLLYSQAGSIYIENKEKTPLOV-AKGLGLILKRWVE 225
Db 190 AILMDHGTDKTLTKGFTPLHIAAKYGNLPVAKSLLE 227

RESULT 12
US-08-847-429A-23
; Sequence 23, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429A
; FILING DATE: 24-APR-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 303 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-847-429A-23

Query Match 22.5%; Score 262; DB 2; Length 303;

Best Local Similarity 33.9%; Pred. No. 7.9e-22;
Matches 74; Conservative 35; Mismatches 97; Indels 12; Gaps 4;
QY 14 AYSKLEELKESILADKSLATRTDQDSR-----TALHWASAGHTEIVEFLQLQGVVND 68
Db 16 AHCGHVRVAK--LLLDNRN---ADPNARALNGFTPLHIACKKNRIKIVELLKYHAAIEA 69
QY 69 KDDAGWSPLHIAASAGRDEIVKALLGKGAQVNAVQNGCTPLHYAASKNRHEIAVMLEGG 128
Db 70 TTESGLSPLHVAAPMGAINIVYLLQQGANADVATVGTETPLHLAARANQTDIVRVLRN 129
QY 129 GANPDADKHYEATAMHRAAAGNLMKMIHILLYKASTNIQDTEGNTPLHLACDEERVEEA 188
Db 130 GAQVDAARAARELQPLHIASRLGNTDIVILLQANASPNAAATRDLYTPLHIAAKEGQEEVA 189
QY 189 KLLYSQAGSIYIENKEKTPLOV-AKGLGLILKRWVE 225
Db 190 AILMDHGTDKTLTKGFTPLHIAAKYGNLPVAKSLLE 227

RESULT 13
US-09-065-474-23
; Sequence 23, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-CI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 303 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-065-474-23

Query Match 22.5%; Score 262; DB 3; Length 303;
Best Local Similarity 33.9%; Pred. No. 7.9e-22;
Matches 74; Conservative 35; Mismatches 97; Indels 12; Gaps 4;

QY 14 AYSKLEELKESILADKSLATRTDQDSR-----TALHWASAGHTEIVEFLQLQGVVND 68
Db 16 AHCGHVRVAK--LLLDNRN---ADPNARALNGFTPLHIACKKNRIKIVELLKYHAAIEA 69
QY 69 KDDAGWSPLHIAASAGRDEIVKALLGKGAQVNAVQNGCTPLHYAASKNRHEIAVMLEGG 128

Db 70 TTESGLSLPLHVAAGFAINIVYLLQGANADVATVGTETPLHLAARANQTDIVRVLVN 129
Qy 129 GANPDADKHYEATAMHRAAAKGNLMIHLLYYKASTNIQDTGNTPLHLACDEERVEA 188
Db 130 GAQVDAARELQTPHLIASRLGNTDIVILLQANASPNAAATRDLYTPLHIAAKEGQEEVA 189
Qy 189 KLLVSGASIIYIENKEEKTPLQV-AKGGGLGLILKRMVE 225
Db 190 AILMDHGTDTKTLTKKGFTPLHLAAKYGNLPVAKSLLE 227

RESULT 14
US-09-188-930-334
; Sequence 334, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Orrust, Rene
; APPLICANT: Murlison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 334
; LENGTH: 787
; TYPE: PRT
; ORGANISM: Mouse
US-09-188-930-334

Query Match 22.2%; Score 258.5; DB 4; Length 787;
Best Local Similarity 29.5%; Pred. No. 8.9e-21;
Matches 71; Conservative 43; Mismatches 110; Indels 17; Gaps 3;
Qy 2 EECVNLNMCVN-----LAYSGKLEELKESILADKSLATRTDQDSRTALHW 46
Db 454 EECVWLLNNANPNLTNRKSGSTPLHMAVERKGRGIVELLARKTSVNAKDEQWTLHF 513
Qy 47 ACSAGHTEIVEELLQGVVNDKDDAGSPLHIAASAGDEIVKALLGKAQVNAVQNG 106
Db 514 AAQNGDEASTRLLEKNASVNEVDEGRTPMHVACQGOENIVRTLRRGVVDVGLQKDA 573
Qy 107 CTPLHYAASKNRHETAVMLL-EGGANPDADKHYEATAMHRAAAKGNLMIHLLYYKAST 165
Db 574 WPLHYAAQGHLPVVKLLAKOPGVSVNAQTLDGRTPLHLAQRHYRVARILIDLCSDV 633
Qy 166 NTQDTGNTPLHLACDEERVEAKLLVSGASIIYIENKEEKTPLQV-AKGGGLGLILKRMV 224
Db 634 NICSLQAQTPHVAAGTGTSTARULLLHRGACKEALTSEGYTALHLAAQNGHLATVKLLI 693
Qy 225 E 225
Db 694 E 694

RESULT 15
US-08-810-712-10
; Sequence 10, Application US/08810712G
; Patent No. 6160106
; GENERAL INFORMATION:
; APPLICANT: Yeda Research and Development Co. LTD
; TITLE OF INVENTION: Tumor Suppressor Genes, Proteins Encoded Thereby and
; TITLE OF INVENTION: Use of said Genes and Proteins
; FILE REFERENCE: sequencelist
; CURRENT APPLICATION NUMBER: US/08/810,712G
; CURRENT FILING DATE: 1997-03-03
; EARLIER APPLICATION NUMBER: PCT/US94/11598

; EARLIER FILING DATE: 1994-10-12
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 1423
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-810-712-10
Query Match 21.2%; Score 247; DB 4; Length 1423;
Best Local Similarity 29.0%; Pred. No. 4.7e-19;
Matches 61; Conservative 33; Mismatches 82; Indels 34; Gaps 1;
Qy 37 QDQSR TALHWACSHAGHTEIVEELLQGVVNDKDDAGSPLHIAASAGRDE----- 87
Db 401 DRGGSNAVYMAARHGHVDTLKFLSENKCPDLVDKSGEMALHVAARYGHADVAQVTCAS 460
Qy 88 -----IVKALLGKAQVNAVQNGCTPLHYAASKNRHETA 122
Db 461 AQIPISRTKEEETPLHCAAHGYYSVAKALCEAGCNVNIKNREGETPLLTASARGYHDIV 520
Qy 123 VMLLEGGANPDADKHYEATAMHRAAAKGNLMIHLLYYKASTNIQDTGNTPLHLACDE 182
Db 521 ECLAEHGAADLNACDKDGHIALHLAVRRQCMQEVIKTLLSQGCFVDYQDRHGNTPLHVACKD 580
Qy 183 ERVEEAKLLVSGASIIYIENKEEKTPLQVA 212
Db 581 GNMPIVVVALCEANCNLDISNKYGRTPHLHA 610
Search completed: August 13, 2001, 07:45:21
Job time: 58 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 2001, 07:44:24 ; Search time 32.54 Seconds
(without alignments)
918.898 Million cell updates/sec

Title: US-09-509-775-2

Perfect score: 1164

Sequence: 1 MEGCVSNLMVCNLAYSGKLE.....TPLQVAKGGIGLILKRWVEG 236

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL16:*

- 1: sp-archaea:*
- 2: sp-bacteria:*
- 3: sp-fungi:*
- 4: sp-human:*
- 5: sp-invertebrate:*
- 6: sp-mammal:*
- 7: sp-mhc:*
- 8: sp-organelle:*
- 9: sp-phage:*
- 10: sp-plant:*
- 11: sp-rodent:*
- 12: sp-unclassified:*
- 13: sp-vertebrate:*
- 14: sp-virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1164	100.0	226	4	075832 homo sapien
2	1106	95.0	231	11	Q922X3
3	1105	94.9	231	11	Q922X2
4	582.5	50.0	118	4	O95533
5	338	29.0	247	10	Q92Q79
6	312.5	26.8	1943	11	Q61307
7	310	26.6	237	3	Q9P3P6
8	310	26.6	1088	4	Q13484
9	310	26.6	4377	4	Q12955
10	306	26.3	2622	11	Q70511
11	305	26.2	636	5	Q9T87
12	301	25.9	1762	11	O88521
13	297	25.5	1159	5	Q9NCP8
14	297	25.5	2443	5	Q9VSA2
15	290.5	25.0	1136	6	Q9N180
16	290	24.9	1719	4	Q13768
17	290	24.9	1856	4	Q99407
18	288.5	24.8	1786	5	Q17344
19	288.5	24.8	1809	5	Q17487

20	288.5	24.8	1815	5	Q17488
21	288.5	24.8	1867	5	Q17486
22	288.5	24.8	2039	5	Q17489
23	288.5	24.8	6994	5	Q17343
24	287.5	24.7	1848	11	Q61302
25	286	24.6	843	11	P97582
26	285.5	24.5	1098	11	Q61304
27	283.5	24.4	1059	4	O15084
28	283.5	24.4	1166	4	O9H2K2
29	283.5	24.4	1265	4	O9HAS4
30	281.5	24.2	1327	4	O95271
31	281	24.1	1181	5	O9XZ37
32	281	24.1	1181	5	O9VBP3
33	276.5	23.8	1549	5	O24241
34	276.5	23.8	1549	5	O9V4B1
35	270.5	23.2	815	6	O9TV77
36	270.5	23.2	1030	4	O14974
37	269.5	23.2	658	11	O62937
38	269.5	23.2	976	11	Q10728
39	268.5	23.1	963	13	Q90624
40	268.5	23.1	1004	13	Q90623
41	266.5	22.9	1395	5	Q9XZC0
42	266	22.9	768	4	Q90LJ7
43	263.5	22.6	2119	5	O9VAU5
44	260.5	22.4	1062	11	O89019
45	260	22.3	1430	11	Q9JJP7

ALIGNMENTS

RESULT 1
O75832 PRELIMINARY; PRT; 226 AA.
AC O75832;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE 26S PROTEASOME SUBUNIT P28.
GN DJ889N15.2 OR HUMAN GANKYRIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98382579; PubMed=9714768;
RA Hori T., Kato S., Saeki M., Demartino G.N., Slaughter C.A.,
RA Takeuchi C., Toh E.A., Tanaka K.;
RT "cDNA cloning and functional analysis of p28 (Nas6p) and p40.5
RT (Nas7p), two novel regulatory subunits of the 26S proteasome.";
RL Gene 216:113-122(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Grafham D.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Higashitsuji H., Fujita J.;
RT "Enhanced expression of a novel tumour marker in the human
RT hepatomas.";
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB009619; BAA33215.1; -;
DR EMBL; AL031177; CAA20117.1; -;
DR HSSP; P42773; 1IHB.
DR InterPro; IPR002110; -;
DR Pfam; PF00023; ank; 5.
DR PROSITE; PS50086; ANK_REPEAT; 5.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR SMART; SM00248; ANK; 1.
KW Proteasome.

Q17488 caenorhabdi
Q17486 caenorhabdi
Q17489 caenorhabdi
Q17343 caenorhabdi
Q61302 mus musculu
P97582 rattus norv
Q61304 mus musculu
O15084 homo sapien
O9H2K2 homo sapien
O9HAS4 homo sapien
O95271 homo sapien
Q9XZ37 drosophila
Q9VBP3 drosophila
O24241 drosophila
O9V4B1 drosophila
Q9TV77 sus scrofa
O14974 homo sapien
O62937 rattus norv
Q10728 rattus norv
Q90624 gallus gall
Q90623 gallus gall
Q9XZC0 latrodectus
Q9UIJ7 homo sapien
Q9VAU5 drosophila
O89019 mus musculu
Q9JJP7 mus musculu

```
SQ SEQUENCE 226 AA; 24428 MW; 57158E331466C7C8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 1164; DB 4; Length 226;
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEGCVSNLMVCNLAYSCKLEELKESILADKSLATRTDQDSRTALHWACSAAGHTEIVEFLL 60
DB 1 MEGCVSNLMVCNLAYSCKLEELKESILADKSLATRTDQDSRTALHWACSAAGHTEIVEFLL 60
QY 61 QLVGVNDKDDAGWSPLHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHE 120
DB 61 QLVGVNDKDDAGWSPLHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHE 120
QY 121 IAVMLLEGGANPDADKHYEATAMHRAAKGNLKMHIHLLYKASTNIQDTGEGNTPPLHLAC 180
DB 121 IAVMLLEGGANPDADKHYEATAMHRAAKGNLKMHIHLLYKASTNIQDTGEGNTPPLHLAC 180
QY 181 DEERVEEAKLLVSQASIIYIENKEEKTPLQVAKGGLGLILKRMVEG 226
DB 181 DEERVEEAKLLVSQASIIYIENKEEKTPLQVAKGGLGLILKRMVEG 226

RESULT 2
ID Q922X3 PRELIMINARY; PRT; 231 AA.
AC Q922X3;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DE GANKYRIN HOMOLOGUE.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Higashitsuji H., Fujita J.;
RT "Cloning of rat gankyrin homologue containing ankyrin repeats.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022014; BAA36954.1; -
DR HSSP; P42773; 1IHB.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank; 5.
DR PROSITE; PS00088; ANK_REPEAT; 5.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 231 AA; 24985 MW; F5241DC9A816066E CRC64;

Query Match
Best Local Similarity 95.08; Score 1106; DB 11; Length 231;
Matches 213; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

QY 1 MEGCVSNLMVCNLAYSCKLEELKESILADKSLATRTDQDSRTALHWACSAAGHTEIVEFLL 60
DB 1 MEGCVSNLMVCNLAYSCKLEELKESILADKSLATRTDQDSRTALHWACSAAGHTEIVEFLL 60
QY 61 QLVGVNDKDDAGWSPLHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHE 120
DB 61 QLVGVNDKDDAGWSPLHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHE 120
QY 121 IAVMLLEGGANPDADKHYEATAMHRAAKGNLKMHIHLLYKASTNIQDTGEGNTPPLHLAC 180
DB 121 IAVMLLEGGANPDADKHYEATAMHRAAKGNLKMHIHLLYKASTNIQDTGEGNTPPLHLAC 180
QY 181 DEERVEEAKLLVSQASIIYIENKEEKTPLQVAKGGLGLILKRMVE 225
DB 181 DEERVEEAKLLVSQASIIYIENKEEKTPLQVAKGGLGLILKRIAE 225

RESULT 3
ID Q922X2 PRELIMINARY; PRT; 231 AA.
AC Q922X2;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE GANKYRIN.
GN PSMD10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Higashitsuji H., Fujita J.;
RT "Cloning of mouse gankyrin containing ankyrin repeats.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022022; BAA36969.1; -
DR HSSP; P42773; 1IHB.
DR MGD; MGI:1858898; Psm10.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank; 5.
DR PROSITE; PS00088; ANK_REPEAT; 5.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 231 AA; 25115 MW; 5A3DB7027B0694E6 CRC64;

Query Match
Best Local Similarity 94.9%; Score 1105; DB 11; Length 231;
Matches 210; Conservative 10; Mismatches 5; Indels 0; Gaps 0;

QY 1 MEGCVSNLMVCNLAYSCKLEELKESILADKSLATRTDQDSRTALHWACSAAGHTEIVEFLL 60
DB 1 MEGCVSNLMVCNLAYSCKLEELKESILADKSLATRTDQDSRTALHWACSAAGHTEIVEFLL 60
QY 61 QLVGVNDKDDAGWSPLHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHE 120
DB 61 QLVGVNDKDDAGWSPLHIAASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHE 120
QY 121 IAVMLLEGGANPDADKHYEATAMHRAAKGNLKMHIHLLYKASTNIQDTGEGNTPPLHLAC 180
DB 121 IAVMLLEGGANPDADKHYEATAMHRAAKGNLKMHIHLLYKASTNIQDTGEGNTPPLHLAC 180
QY 181 DEERVEEAKLLVSQASIIYIENKEEKTPLQVAKGGLGLILKRMVE 225
DB 181 DEERVEEAKLLVSQASIIYIENKEEKTPLQVAKGGLGLILKRIAE 225

RESULT 4
ID Q95533 PRELIMINARY; PRT; 118 AA.
AC Q95533;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE DJ889N15.2.2 (26S PROTEASOME SUBUNIT P28 (ANKYRIN REPEAT PROTEIN))
DE (PUTATIVE PARTIAL ISOFORM 2) (FRAGMENT).
GN DJ889N15.2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Grafham D.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL031177; CAA20118.1; -
DR HSSP; Q13625; 1VCS.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank; 2.
```

```

Qy 192 VSGASITYENKEKPTLPQAVGGGLGIL- ----KRWEG 226
   : || :||| | | | | | | | | | | | | | |
Db 209 IRHGADVDEKGYTVLGRATNEFRPALIDAAKAMLEG 247

RESULT 6
ID Q61307 PRELIMINARY; PRT: 1943 AA.
AC Q61307; Q61306; Q61308; Q61309; Q61310; Q08866; Q08867;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ANKYRIN 3 (ANKYRIN G) (EPITHELIAL ANKYRIN) (ANKYRIN-3).
GN ANK3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A. (ISOFORMS 1-6).
RP STRAIN=C57BL/6J; TISSUE=KIDNEY;
RC MEDLINE=95340633; PubMed=9060470;
RA Peters L.C., John K.M., Lu F.M., Eicher E.M., Yialamas M.,
RA Turtzo L.C., Otsuka A.J., Lux S.E.;
RT "Ank3 (epithelial ankryn), a widely distributed new member of the
RT ankryn gene family and the major ankryn in kidney, is expressed in
RT alternatively spliced forms, including forms that lack the repeat
RT domain.";
RL J. Cell Biol. 130:313-330(1995).
RN [2]
SEQUENCE FROM N.A. (ISOFORMS 7 AND 8).
RP STRAIN=C3H/HEJ; TISSUE=BONE MARROW;
RC MEDLINE=97213781; PubMed=9060470;
RA Hoock T.C., Peters L.L., Lux S.E.;
RT "Isoforms of ankryn-3 that lack the NH2-terminal repeats associate
RT with mouse macrophage lysosomes.";
RL J. Cell Biol. 136:1059-1070(1997).
RN [-]
FUNCTIONAL ELEMENTS; THEY BIND TO THE INTEGRAL MEMBRANE PROTEINS TO CYTO-
SC SKELETAL ELEMENTS; THEY BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN
BAND 4.2, TO NA-K ATPASE, TO THE LYMPHOCTE MEMBRANE PROTEIN GP85,
AND TO THE CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND
DESMIN. ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO
THE CYTOSOLIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;
THEY REPAIR MOST OR ALL OF THESE BINDING FUNCTIONS.
CC -1 SUBCELLULAR LOCATION: PLASMA-MEMBRANE ASSOCIATED. ISOFORMS 7 AND 8
CC ARE ASSOCIATED WITH THE LYSOSOMAL MEMBRANE.
CC -1 ALTERNATIVE PRODUCTS: AT LEAST 8 ISOFORMS; 1 (SHOWN HERE), 2, 3,
CC 4, 5, 6, 7 AND 8; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1 TISSUE SPECIFICITY: MAJOR FORM OF ANKYRIN IN KIDNEY AND OTHER
CC EPITHELIAL TISSUES. ALSO EXPRESSED IN MACROPHAGES, MEGAKARYOCYTES,
CC LEYDIG CELLS, CARDIAC, SMOOTH AND SKELETAL MUSCLE, INITIAL
CC SEGMENTS OF AXONS, AND NODES OF RANVIER.
CC -1 SIMILARITY: CONTAINS 24 ANK REPEATS.
DR ENBL; L40631; AAB01603.1; -
DR ENBL; L40632; AAB01605.1; -
DR ENBL; U89275; AAB58381.1; -
DR ENBL; L40631; AAB01602.1; -
DR ENBL; L40632; AAB01604.1; -
DR ENBL; L40632; AAB01606.1; -
DR ENBL; L40632; AAB01607.1; -
DR ENBL; U89274; AAB58380.1; -
DR HSP; P55273; IBD8.
DR MGD; MG1:88026; Ank3.
DR InterPro; IPR000488; -.
DR InterPro; IPR000906; -.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; z05; 1.
DR PROSITE; PSS0088; ANK_REPEAT; 21.
DR PROSITE; PSS0297; ANK_REPEAT_REGION; 1.
DR PROSITE; PSS0017; DEATH_DOMAIN; 1.
DR SMART; SW00005; DEATH; 1.

```

KW Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
KW Multigene family.
KW DOMAIN 1 826 ANION EXCHANGE PROTEIN BINDING DOMAIN.
FT DOMAIN 23 810 24 X ANK MOTIF REPEATS.
FT DOMAIN 856 1437 SPECTRIN BINDING DOMAIN.
FT DOMAIN 1438 1943 REGULATORY DOMAIN (REGULATES BINDING OF
ANKYRIN TO SPECTRIN AND THE BAND 3
PROTEIN).
FT REPEAT 23 55 ANK MOTIF 1.
FT REPEAT 56 88 ANK MOTIF 2.
FT REPEAT 89 121 ANK MOTIF 3.
FT REPEAT 122 154 ANK MOTIF 4.
FT REPEAT 155 183 ANK MOTIF 5.
FT REPEAT 184 216 ANK MOTIF 6.
FT REPEAT 217 249 ANK MOTIF 7.
FT REPEAT 250 282 ANK MOTIF 8.
FT REPEAT 283 315 ANK MOTIF 9.
FT REPEAT 316 348 ANK MOTIF 10.
FT REPEAT 349 381 ANK MOTIF 11.
FT REPEAT 382 414 ANK MOTIF 12.
FT REPEAT 415 447 ANK MOTIF 13.
FT REPEAT 448 480 ANK MOTIF 14.
FT REPEAT 481 513 ANK MOTIF 15.
FT REPEAT 514 546 ANK MOTIF 16.
FT REPEAT 547 579 ANK MOTIF 17.
FT REPEAT 580 612 ANK MOTIF 18.
FT REPEAT 613 645 ANK MOTIF 19.
FT REPEAT 646 678 ANK MOTIF 20.
FT REPEAT 679 711 ANK MOTIF 21.
FT REPEAT 712 744 ANK MOTIF 22.
FT REPEAT 745 777 ANK MOTIF 23.
FT REPEAT 778 810 ANK MOTIF 24.
FT VARSPLIC 1 805 MISSING (IN ISOFORM 7 AND ISOFORM 8).
FT VARSPLIC 1 849 MISSING (IN ISOFORM 3).
FT VARSPLIC 833 853 MISSING (IN ISOFORM 7 AND ISOFORM 8).
FT VARSPLIC 834 855 VKASAPKLSGDGIYSIDGEEG -> GDKCTWFKIPKQVEV
LVK (IN ISOFORM 5).
FT VARSPLIC 850 855 SDGEEG -> MALPHS (IN ISOFORM 2 AND
ISOFORM 3).
FT VARSPLIC 852 852 G -> GDKCTWFKIPKQVEVLVKS (IN ISOFORM 4
AND ISOFORM 6).
FT VARSPLIC 1569 1764 MISSING (IN ISOFORM 3, ISOFORM 4 AND
ISOFORM 7).
SQ SEQUENCE 1943 AA; 212013 MW; B23C4B6F82904804 CRC64;
Query Match 26.8%; Score 312.5; DB 11; Length 1943;
Best Local Similarity 35.5%; Pred. No. 5e-19; Mismatches 67; Indels 33; Gaps 1;
Matches 72; Conservative 31; Mismatches 67; Indels 33; Gaps 1;
QY 42 TALHWACSAGHTEIVEFLLQLGVPVNDKDDAGWSPLHTAASAGRDEIVKALLGK----- 95
DB 517 TPLHLAAREGHEDVAFLLDHGLSITTKGFTPLHVAAYKGLVSLLLQKSASPDA 576
QY 96 -----GAQVNAVNGCTPLHYAASKNRHEIAVMLEGG 128
DB 577 AGKSGLTPLHVAHYDNDKVALLLDQAGSPHAAKNGYTPLHTAAKNQMDIATSLLEY 636
QY 129 GANPKADHYEATAMHRAAAGNLKMIHLLYYKASTNIQTEGNTPLHLACDEERVEEA 188
DB 637 GADANAVTQGIASVHLAAGQGVDMVSLLSRNANVLSNKGSLTPUHLAAQEDRVNVA 696
QY 189 KLLVSGASIIYENKEETPLQV 211
DB 697 EVLVNQAQHVDAQTGMGTPLHV 719
RESULT 7
QY9P3P6 PRELIMINARY; PRT; 237 AA.
AC Q9P3P6;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)

DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE RELATED TO 26S PROTEASOME SUBUNIT P28.
GN B7F18.30.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Farlmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL389891; CAB97304.1; -
DR InterPro; IPR002110; -
DR Pfam; PF00023; ank; 6.
DR PROSITE; PS50088; ANK_REPEAT; 3.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR SMART; SM00248; ANK; 1.
KW Proteasome.
SQ SEQUENCE 237 AA; 25995 MW; 6E2B6EFC69D909B1 CRC64;
Query Match 26.6%; Score 310; DB 3; Length 237;
Best Local Similarity 34.5%; Pred. No. 4.9e-20;
Matches 78; Conservative 30; Mismatches 90; Indels 28; Gaps 4;
QY 14 AYSKLEELKESILADSLATRTDQDSRTALHWACSAGHTEIVEFLL-QLGVPVNDKDDA 72
DB 12 ARDGKASIVESLLNANPKLAQRKDDGRLPLTHWACSYNRKEVELLVNQKGFDPDVEDDM 71
QY 73 GWSPLHTAASA-GDEIVKALLGKAQVNAVNGCTPLHYAASKNRHEIAVMLEGGAN 131
DB 72 GWTFFMISASVKSDSAIDLLSLSGADINOTNHOSQALHFIAKSNIDILARKLLSPDMK 131
QY 132 PD-----AKDHYEATAMHRAAAGNLKMIHLLYYKASTNIQTEGNTPLHLACDEERVE 186
DB 132 PKPASVRVKRGQYPLHRAAAGVPMINLLQHKSPINASDNAGVTPHLHVAEGHGH 191
QY 187 EAKLLVSGQASI-----YIENKEETPLQV 211
DB 192 AAVALLKAGAEATDKKMDGYLALDLPADPKVRRFIEKEAKEGIEL 237
RESULT 8
QY13484 PRELIMINARY; PRT; 1088 AA.
AC Q13484;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE ANKYRIN G119.
GN ANK3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RX MEDLINE=96234670; Pubmed=8666667;
RA Kashgarian M., Morrow J.S., Ardito T., Stabach P.R., Mann A.S.,
RA Devorajan P.;
RT "Identification of a small cytoplasmic ankyrin (AnkG119) in the kidney
and muscle that binds beta 1 sigma spectrin and associates with the
Golgi apparatus."
RL J. Cell Biol. 133:819-830(1996).
DR EMBL: U43965; AAB08437.1; -
DR HSP; P55273; 1BD8.
DR InterPro; IPR000906; -

Query Match 26.6%; Score 310; DB 4; Length 4377;
Best Local Similarity 37.0%; Pred. No. 2.5e-18;

```
Db 553 EDVAAFLLDHGASLITTKGFTPLHVAAYKGLKLEVASLLQLKSAPDAAGKSGLTPLHV 612
QY 80 AASAGRDEIVKALLGKAQVNAVNGCTPLHYAASKNRHEIAVMLLEGANPDADKHYE 139
Db 613 AAHYDNQKVALLLDQOGASPHAAKNGYTPPLHYAAKKNQMDIATSLLEYGADANPYTROG 672
QY 140 ATAMHRAAAKGNLKMTHILYYKASTNIQDTEGTPHLHACDEERVEEAKLLYSOGASIY 199
Db 673 IASVHLAAQEGHVDVMSLLSRANVNLNKSGLTPLHLAAQEDRVNAEVLVNOGAHVD 732
QY 200 IENKEEKTPLQV 211
Db 733 AQTMGYTPPLHV 744

RESULT 11
Q9T267 PRELIMINARY; PRT; 636 AA.
AC Q9T267;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE F40G9.1 PROTEIN.
GN F40G9.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Berkson A.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kersey J., Kirsten J., Laister N., Latreille P.,
RA Lighning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Shaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RN Nature 368:32-38(1994).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Graves T., Sutterer C., Ozersky P.;
RT "The sequence of C. elegans cosmid F40G9.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF099919; AAC68798.1; -.
DR HSSP; Q00421; LAWG.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank; 5.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT; 1.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 636 AA; 73229 MW; AB55162AF5D5B1C0 CRC64;
```

```
Query Match 26.2%; Score 305; DB 5; Length 636;
Best Local Similarity 29.3%; Pred. No. 5.2e-19;
Matches 86; Conservative 39; Mismatches 79; Indels 90; Gaps 6;

QY 5 VSNLMVCN-LAYSGLKEELKESILADSLATRTDQDSRTALHWACSGHTEIVEIFLQLG 63
Db 335 ILHKMKCILHYFTKNVYAKRLTRYKLVGYTDDSGRTIHFAVVGSLPLIQFAI--- 391
```

```
QY 64 VPVNDKDDA----- 72
Db 392 --LNDPEMAHKTDVVRVGMNLTKEIENLFENRKICDERKNYSKNTKNPHSEFFETID 449
QY 73 -----GWSPLHIAASAGRDEIVKALLG-KGAQVNA 101
Db 450 FEHENSQNAKKKFFSLKIDIFDLNCLILPLGWTPLMIASSAGRVVRYLLTLPDQDVKH 509
QY 102 VNONGCTPLHYAASKNRHEIAVMLLEGAN-PDAKHYTEATAMHRAAAGNLMKMIHILY 160
Db 510 TNSNKOTSILHYACSKNHVEIVKLLIEADPNINLPDKFGATALHRAASRGNDVIVRALVS 569
QY 161 Y-KASTNIQDTEGTPHLHACDEERVEEAKLLYSOGASIYIENKEEKTPLQVAK 213
Db 570 TGKCSLDRDQGEGETALHACDENRGDVAILLVNRGADMKMLNKEKQTPLEMLK 623

RESULT 12
O88521 PRELIMINARY; PRT; 1762 AA.
AC O88521;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DE 190 KDA ANKYRIN ISOFORM.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98395113; PubMed=9727010;
RA Thevananthar S., Kollu A.H., Devatajan P.;
RT "Identification of a novel ankyrin isoform (AnkG190) in kidney and
RT lung that associates with the plasma membrane and binds alpha-Na, K-
RT ATPase.";
RL J. Biol. Chem. 273:23952-23958(1998).
DR EMBL; AF069525; AAC34809.1; -.
DR HSSP; P55273; 1BI8.
DR InterPro; IPR000169; -.
DR InterPro; IPR000488; -.
DR InterPro; IPR000906; -.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REPEAT; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS00139; THIOLEPROTEASE_CYS; UNKNOWN_1.
DR SMART; SM00005; DEATH; 1.
SQ SEQUENCE 1762 AA; 191898 MW; 9023280086A7BF4E CRC64;
```

```
Query Match 25.9%; Score 301; DB 11; Length 1762;
Best Local Similarity 35.9%; Pred. No. 4.7e-18;
Matches 69; Conservative 36; Mismatches 87; Indels 0; Gaps 0;
```

```
QY 20 EELKESILADKSLATRTDQDSRTALHWACSGHTEIVEIFLQLGVPVNDKDDAGWSPLHI 79
Db 539 EDVAAFLLDHGASLITTKGFTPLHVAAYKGLKLEVASLLQLKSAPDAAGKSGLTPLHV 598
QY 80 AASAGRDEIVKALLGKAQVNAVNGCTPLHYAASKNRHEIAVMLLEGANPDADKHYE 139
Db 599 AAHYDNQKVALLLDQOGASPHAAKNGYTPPLHYAAKKNQMDIATSLLEYGADANAVTROG 658
QY 140 ATAMHRAAAKGNLKMTHILYYKASTNIQDTEGTPHLHACDEERVEEAKLLYSOGASIY 199
Db 659 IASVHLAAQEGHVDVMSLLSRANVNLNKSGLTPLHLGQGEDRVNVAEVLVNOGAHVD 718
QY 200 IENKEEKTPLQV 211
```

```
Db 719 AOTRMGYTPLHV 730
RESULT 13
Q9VSA2 PRELIMINARY; PRT; 1159 AA.
AC Q9VSA2
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-MAR-2001 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ANKYRIN 2
GN ANK2 OR CG7462.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20304926; PubMed=10844021;
RA Bouley M., Tian M.-Z., Paisley K., Shen Y.-C., Malhotra J.D.,
RA Hortsch M.;
RT "The IL-type cell adhesion molecule neuroglian influences the
RT stability of neural ankyrin in the Drosophila embryo but not its
RT axonal localization."
RL J. Neurosci. 20:4515-4523(2000).
DR EMBL: AF190635; AAF73309.1; -
DR FlyBase; FBgn0017645; Ank2.
DR InterPro; IPR000906; -
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00791; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 21.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 1159 AA; 125769 MW; 9ED146C9E483DE58 CRC64;

Query Match 25.5%; Score 297; DB 5; Length 1159;
Best Local Similarity 29.7%; Pred. No. 6.1e-18;
Matches 71; Conservative 43; Mismatches 87; Indels 38; Gaps 4;

Qy 25 SILADK--SLATRTDQDSRTALHWACSAAGHTEIVEFLQLQGVVNDKDGWSP----- 76
Db 255 SLLLEKGGNIEAKT-RDGLTFLHCAARSGEQVDMLLERGAPISAKTKNGLAPLHMAAQ 313
Qy 77 -----LHIAASAGRDEIVKALLGKGAOVNAVNGGCTP 109
Db 314 GEHVDAARILLYHRAPVDEVTVDTLALHVAACHGVVRVAKLLLDNRNADANARALNGFTP 373
Qy 110 LHAAKSKNRHEIAVMLLEGGANPDADKHYEATAMHRAAKAGNLMHIILLYKASTNIQD 169
Db 374 LHIACKKRLKVVLELLRHGASISATTESGLTPLHVAAFMGCMNIVLYLQHDASPDVPT 433
Qy 170 TEGNTPLHLACDEERVEAKLLVSOGASIIYIENKEETPLOVAK--GGIGLILKRMVSG 226
Db 434 VRGETPLHLARANQTDIIRILLRNGAODARAREQQTPLHIASRLGNVDIVMLLLQHG 492

RESULT 14
Q9VSA2 PRELIMINARY; PRT; 2443 AA.
AC Q9VSA2
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAR-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ESTS:149B105 PROTEIN.
GN ANK2 OR ESTS:149B105 OR CG7462.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkov D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003558; AAF50525.1; -
DR HSSP; P42773; 11HB.
DR FlyBase; FBgn0017645; Ank2.
DR InterPro; IPR000488; -
DR InterPro; IPR000906; -
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR SMART; SM00248; ANK; 1.
SQ SEQUENCE 2443 AA; 269538 MW; 7561DC5CF56812DC CRC64;

Query Match 25.5%; Score 297; DB 5; Length 2443;
Best Local Similarity 29.7%; Pred. No. 1.7e-17;
Matches 71; Conservative 43; Mismatches 87; Indels 38; Gaps 4;

Qy 25 SILADK--SLATRTDQDSRTALHWACSAAGHTEIVEFLQLQGVVNDKDGWSP----- 76
Db 255 SLLLEKGGNIEAKT-RDGLTFLHCAARSGEQVDMLLERGAPISAKTKNGLAPLHMAAQ 313
Qy 77 -----LHIAASAGRDEIVKALLGKGAOVNAVNGGCTP 109
Db 314 GEHVDAARILLYHRAPVDEVTVDTLALHVAACHGVVRVAKLLLDNRNADANARALNGFTP 373
Qy 110 LHAAKSKNRHEIAVMLLEGGANPDADKHYEATAMHRAAKAGNLMHIILLYKASTNIQD 169
Db 374 LHIACKKRLKVVLELLRHGASISATTESGLTPLHVAAFMGCMNIVLYLQHDASPDVPT 433
```

QY 170 TCGTPLHLACDEERVEAEKLLVSGQASIIYENKEKTPLOVAK--GGLGLILKRMVEG 226
| ||||| : : : : : || : : : : : || : : : : :
Db 434 VRGETPLHLAARANQTDIIIRILLRNGAQVDARAREQOTPLHLIASRLGNVDIVMLLQHG 492

RESULT 15

Q9N180
ID Q9N180 PRELIMINARY; PRT: 1136 AA.
AC Q9N180;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ANKYRIN 1 (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_taxid=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRED JAPANESE BLACK; TISSUE=BONE MARROW;
RA Matsumoto M., Inaba M., Koshino I., Saito D., Ono K.;
RT "Cloning of bovine erythrocyte ankyrin";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF222766; AAF61702.1; -
DR InterPro: IPR000906; -
DR InterPro: IPR002110; -
DR Pfam: PF00023; ank; 24.
DR Pfam: PF00791; ZUS; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REGION; 1.
DR SMART; SM00248; ANK; 1.
FT NON_TER 1136 1136
SQ SEQUENCE 1136 AA; 123182 MW; 07CD8B1568ACE8C7 CRC64;

Query Match 25.0%; Score 290.5; DB 6; Length 1136;
Best Local Similarity 37.5%; Pred. No. 2.3e-17;
Matches 72; Conservative 30; Mismatches 81; Indels 9; Gaps 2;

QY 26 ILADKSLATRTDODSR-----TALHWACSAGHTEIVEFLQLGVNPNDRKDDAGWSPLHIA 80
Db 355 VLLDKG-----AKPNSRALNGFTPLHIACKNHRVMEILLKMGASIDAVTESGLTPLHVA 410
QY 81 ASAGRDEIVKALLGKGAQVNAVNGCTPLHYAASKNRHEIAVMLEGGANPDAKDHYEA 140
Db 411 SEFGHPPIVKSLLQREASPNVSNVETPLHMAARAGHTEVAKYLLQNKAKVNAKAKDDQ 470
QY 141 TAMHRAAKGNLKMHIILLYKASTNIQDTGNTPLHLACDEERVEAEKLLVSGQASIIYI 200
Db 471 TPLHCAARIGHTNNVKILLENNANPNLATAGHTPLHIAAREGHVETALALLEKEASQTC 530
QY 201 ENKEEKTPLQVA 212
Db 531 MTRKGFTPLHVA 542

Search completed: August 13, 2001, 07:47:24
Job time: 180 sec